

THE NETWORK STRUCTURE OF INTERNATIONAL TRADE*

Thomas CHANEY[†]

University of Chicago, NBER and CEPR

November, 2011

Abstract

I build a dynamic model of the formation of an international network of importers and exporters. Firms can only export into markets in which they have a contact. They directly search for new trading partners, but they can also use their existing network of contacts to remotely search for new partners. This model explains *(i)* the cross-sectional distribution of the number of foreign markets accessed by individual exporters, *(ii)* the cross-sectional geographic distribution of foreign contacts, and *(iii)* the dynamics of firm level exports. I show that the firm level dynamics of trade can explain the observed cross section of firm level exports. All theoretical predictions are supported by the data.

Introduction

Individual firms differ hugely in their exposure to international trade. Most firms do not export abroad. Of those which do, only few export to a large number of countries. This heterogeneity in the access to foreign markets of individual firms has dramatic implications for the patterns of international trade. Melitz (2003) shows that, in the presence of heterogeneity in the ability of individual firms to access foreign markets, a reduction in trade barriers can induce aggregate productivity gains. Bernard, Eaton, Jensen and Kortum (2003) and Chaney (2008) show that with firm heterogeneity, firm level exports aggregate up to the well established gravity equation in international trade, but that the sensitivity of trade flows with respect to trade barriers is magnified. The source of this heterogeneity in the ability of individual firms to access foreign markets however remains largely unexplained. Whereas Bernard, Eaton, Jensen and Kortum

*I am grateful to Enghin Atalay, Sylvain Chassang, Xavier Gabaix, Sam Kortum, Pierre-Louis Lions, Bob Lucas, Marc Melitz, Roger Myerson, David Sraer, and seminar participants at Chicago (Math and Econ), Columbia, Harvard, MIT, the NBER Summer Institute, NYU, Princeton, Sciences Po (Paris), Toronto, the Toulouse School of Economics, UQAM, UW Milwaukee, Wharton and Yale for helpful discussions, and NSF grant SES-1061622 for financial support. I am indebted to Ferdinando Monte and Enghin Atalay for their superb research assistance.

[†]Contact: Department of Economics, The University of Chicago, Chicago, IL 60637. Tel: 773-702-5403. Email: tchaney@uchicago.edu.

(2003) or Melitz (2003) assume that this heterogeneity is entirely driven by productivity differences, Armenter and Koren (2009) point out that productivity differences can only account for a fraction of the exposure to international markets.

In this paper, I offer a simple explanation for the heterogeneous ability of individual firms to access foreign markets based on the formation of an international network of exporters and importers. Individual exporters meet foreign importers both via direct search, and through their network of existing foreign contacts via remote search. The cross-sectional predictions of the model on both the distribution of the number of foreign contacts, and on the geographic distribution of foreign contacts are supported by the data. Furthermore, this model generates novel predictions for the dynamic evolution of trade flows. Exporters follow a history dependent path when expanding into foreign markets. The entry of individual exporters into a given country is influenced by changes in aggregate trade flows between third countries, in a way that is consistent with the model and with the cross-sectional evidence on the distribution of foreign contacts.

Potential exporters meet foreign trading partners in two distinct ways. First, a firm directly searches for foreign partners, which I model as a geographically biased random search. This is a reduced form for the active search for foreign trading partners. Second, once a firm has acquired a network of foreign contacts in various foreign locations, it can remotely search for new trading partners from these locations. The possibility to use existing contacts to find new ones gives an advantage to firms with many contacts. This generates a fat tailed distribution for the number of foreign contacts across firms. The key parameter that shapes the cross-sectional distribution of the number of contacts is the relative efficiency of direct and remote search. The empirical distribution of the number of foreign contacts is well described by the theoretical model. This allows me to estimate precisely the relative efficiency of direct and remote search. Moreover, direct evidence on the time-series evolution of firm level trade flows confirms the assumed mechanism. I find that the more contacts a firm has, the more likely it is to acquire additional contacts. I tie together quantitatively the cross-sectional and time-series evidence on firm level trade.

The more novel contribution of this paper is that the dynamic formation of a network of trading partners is embedded into geographic space. Remote search allows say a French exporter that has acquired a contact in Japan to radiate away from Japan as Japanese firms would. It does so by using its Japanese contacts as a remote hub from which it can expand out of Japan. The theory therefore predicts that as firms acquire more foreign contacts, they expand into more remote

countries, so that their exports become geographically more distant. The speed at which the geographic distance of exports increases depends on the relative importance of direct and remote search. I show empirically that the geographic distance of exports increases with the number of foreign contacts in a way that is quantitatively in line with the theory and the cross-sectional distribution of the number of foreign contacts.

This is a theory of a network. Therefore, a shock that hits anywhere will be transmitted to all the components in the network, with an intensity that depends on the structure of the network. I find empirical support for these novel predictions on the dynamics of firm level trade flows. For instance, I show that an increase in the volume of trade between country a and b will have a positive impact on the probability that a French firm that already exports to a starts exporting to b , but not on firms that do not export to country a yet. The magnitude of this effect is in line qualitatively and quantitatively with the theory and the cross-sectional distribution of the number of foreign contacts.

This paper contributes to the literature on both international trade and social networks.

There is a nascent literature in international trade and macroeconomics on the role that social networks and informational barriers play in facilitating or hampering transactions, and in transmitting shocks. In a seminal paper, Rauch (1999) conjectures that informational barriers play an important role in hampering trade. He offers a classification of traded goods between differentiated and homogeneous goods, and shows that geographic proximity is more important for trade in differentiated goods. He argues that this is evidence for the importance of informational barriers. While the Rauch classification has been widely used in international trade, the notion that informational networks are important in overcoming informational barriers has remained relatively under explored. I offer a formal treatment of the network that allows information to diffuse, and show strong evidence of this network using firm level trade data. Rauch and Trindade (2002) show that the presence of ethnic Chinese networks facilitates bilateral trade, and particularly so for trade in differentiated goods compared to trade in homogeneous goods. They argue that these findings are evidence for the importance of informational barriers, and that social networks mitigate those barriers. Rauch (2001) offers a survey of the literature on networks in international trade. In the context of intra-national trade, Combes, Lafourcade and Mayer (2005) show that social and business networks facilitate trade between regions within France, where they use migrations and multi-plant firms to infer a measure of social and business linkages respectively. Using a similar ap-

proach for Spain, Garmendia, Llano, Minondo and Requena (2011) show that social and business networks have a stronger impact on the extensive margin than on the intensive margin of trade. Burchardi and Hassan (2010) show that West German regions that have closer social ties with East Germany experienced faster growth and engaged in more investment into East Germany after the German reunification. On a somewhat related topic, Hidalgo, Klinger, Barabási and Hausmann (2007) show that the product mix of goods manufactured and exported by countries can be described as a network, and that countries move towards more connected sectors as they grow. Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi (2011) describe the input-output linkages between sectors in the U.S. as a network, and show how idiosyncratic shocks to individual sectors have a non negligible impact on aggregate volatility. In this paper, I develop a more general model of the formation of an international network of firms, and show how this network matters for firm level trade patterns, over and beyond the effects analyzed in the relatively narrow special cases studied so far.

This paper is also related to a recent literature that emphasises the importance of informational barriers in international trade and the role of trade intermediaries in overcoming those barriers. Casella and Rauch (2002) offer a formal model of trade with informational barriers. They assume that there are only two types of agents, some are perfectly informed about the quality of foreign goods, while the others are uninformed. The informed agents may chose to act as intermediaries for international trade. I offer a more nuanced model where firms gradually learn about foreign markets, so that there is close to a continuum of firms with a differential access to information about foreign markets. Antràs and Costinot (2011) develop a theoretical model of trade that relaxes the assumption of a centralized Walrasian market, and derive predictions for the welfare gains from trade in a setting where trade is intermediated. Ahn, Khandelwal and Wei (2011) demonstrate empirically the importance of trade intermediaries in facilitating trade, especially for smaller exporters and for penetrating less accessible markets. I do not formally introduce trade intermediaries, but I stress the importance of informational barriers, and show how a network can partially overcome these barriers. The network I describe can be thought of as a formal treatment of the way in which trade intermediaries connect importers and exporters.

This paper is complementary to models of international trade with heterogeneous firms such as Bernard, Eaton, Jensen and Kortum (2003), Melitz (2003) and its extension in Chaney (2008). Those models assume that differences in the ability of individual firms to enter foreign markets are entirely driven by some exogenous productivity differences, and by the configuration of exogenous

parameters that govern the accessibility of different foreign markets. These models successfully replicate a series of stylized facts regarding the size distribution of individual firms in different markets and the efficiency of firms entering different sets of countries, as shown by Eaton, Kortum and Kramarz (2011). But they take as exogenous all the parameters that govern the entry of firms into different markets. In other words, these models are successful at explaining the intensive margin of trade at the firm level, but are silent about the determinants of the extensive margin of trade. Moreover, these models are unable to match simultaneously the different stylized facts I uncover regarding the distribution of the number and the geographic location of foreign markets entered by different firms. By contrast, the model I develop offers a parsimonious explanation for the extensive margin of trade at the firm level, but is mostly silent about the intensive margin of trade. Because this model is analytically tractable, it would be easy to combine it with a Melitz-type model, and explain both the extensive and intensive margins of trade. In that sense, this model is complementary to the existing models of trade with heterogeneous firms.

This paper is also complementary to a recent literature on the dynamics of exports or more generally expansion at the firm level. Alborno, Calvo Pardo, Corcos and Ornelas (2010) and Defever, Heid and Larch (2010) both present simple models of learning about a firm's potential in a foreign market. They show evidence of the sequential entry respectively of Argentine and Chinese exporters into foreign markets, meaning that where a firm already exports influences where it enters next. Morales, Sheu and Zahler (2011) use a moment inequality estimation procedure to estimate a similar model of sequential export choice, and document that exports tend to be history dependent. They stress the importance of what they call "extended gravity", which is the fact that if a firm exports to a particular country, it is subsequently more likely to export to other similar countries. This corresponds to the notion of remote search in my model. In the case study of a single firm, Jia (2008) and Holmes (2011) study the geographic expansion of Wal-mart in the US. Both stress the importance of local complementarities. New Wal-mart outlets tend to benefit from the proximity of its existing retail centers. While the type of local complementary are similar to the notion of remote search in my paper, they seem to operate in the case of Wal-mart at a much finer geographic level than in the case of exporters. But the typical expansion of this single firm is similar to the expansion of exporters in my model. My paper is complementary to those papers, in the sense that I incorporate these observations formally into a theoretical model of the dynamics of entry of firms, and show how they shape both the cross-sectional distribution of exports as well as the time series of exports at the firm level.

Finally, this paper is related to the literature on social networks. Jackson and Rogers (2007) propose a tractable way to combine the features of a random network and a preferential network.¹ The notions of direct and remote search in my model are similar to their notion of random and preferential attachment. The main theoretical innovation of my model is to embed this general network into a continuous space. For the purpose of this paper, I assume that this space corresponds to the physical geographic space. But it could alternatively correspond to any other space that describes some of the attributes of the agents connected through that network. Recent models of social networks characterize the links formed by agents that differ according to a discrete set of attributes.² Bramoullé and Rogers (2010) consider a model with two types that are biased against each other, while Currarini, Jackson and Pin (2011) consider an extension with more than two types. Both papers show that over time, agents increase the diversity of their contacts, in the sense that they get connected with different types. They derive conditions under which bias eventually vanishes. As the notion of a bias between types is similar to the notion of geographic distance between firms in my model, their results are similar to the gradual geographic expansion of exports in my model. The discrete technique used in those papers is complementary to the continuous approach I use: while I can model a continuum of types, I have to impose an assumption of symmetry that these models relax in the arguably simpler set-up of a finite number of types. I also offer an empirical application of a network model to a data-set much larger than has typically been used in the social network literature.

The remainder of the paper is organized as follows. In section 1, I present a simple theoretical model of the formation of an international network of importers and exporters. In section 2, I test empirically the main theoretical predictions of the model. I relegate to the Appendix all mathematical proofs (Appendix A), and the description of the data and robustness checks (Appendix B).

1 A dynamic model of exports

In this section, I develop a model of the sequential entry of a firm into foreign markets. Over time, this firm acquires a network of customers in various foreign countries, and exports into each

¹See Erdős and Rényi (1959) for a seminal description of random networks, and Barabási and Albert (1999) for a description of preferential networks that exhibit scale-free degree distributions.

²See McPherson, Smith-Lovin and Cook (2001) for an overview of various situations where agents tend to connect to each other according to some attributes outside of the network, which is generally described as homophily.

country where it has some customers. Each period, this firm will both search for new customers at random, and use its existing network of foreign customers to search remotely. In other words, if a French exporter initially enters say Japan, it is subsequently more likely to enter a country with close links to Japan, say Korea. Over time the network of customers of a firm will become more complex: the firm enters foreign countries first at random, and then increasingly through the history dependent process of radiating from existing foreign markets.

1.1 Set-up

Consider the following extension of the Krugman (1980) model with matching frictions.

There is a continuum of consumers, uniformly distributed over the real line \mathbb{R} . A point $x \in \mathbb{R}$ corresponds to a location. For concreteness, one can think of such a location x as a city. A country will then be a collection of cities, or an interval of the real line. This however will be a model of trade between cities, and national boundaries are simply arbitrary lines in the sand. I will introduce formally national boundaries between countries only when I bring the model to the data in Section 2. Time is discrete.

All consumers have the same iso-elastic preferences over a continuum of differentiated goods. All firms face the same increasing returns to scale technology, characterized by a fixed cost of entry, and a constant marginal cost of production. In this simple set-up, all firms charge the same constant mark-up over marginal cost to any consumer that has access to their good. The amount sold to a given consumer depends on the income of that consumer, and on which other alternative goods this consumer has access to. Since I am only interested in the extensive margin of trade, I will not explicitly solve for the equilibrium level of sales of each firm to each of its consumer.³

I assume that the number of firms in any location grows at a constant rate γ . Imposing a free entry condition, characterizing the entry decision of firms, and assuming that the labor force grows at a constant rate γ in every location, one could solve for the steady state growth path of this economy. Instead of solving for this steady state growth path, I directly assume that the population of firms grows at the constant rate γ .

In the absence of any friction, all firms would sell to all consumers. I depart from the Krugman (1980) model and assume that firms face matching frictions. A firm does not have access to all consumer, but instead can only sell to a subset of them. The matching frictions are as follows.

³I develop a simple version of this model in the online Appendix.

See http://home.uchicago.edu/tchaney/research/TradeNetworks_Appendix.pdf

Every period, a firm located in the origin acquires new consumers in two distinct ways. First, the firm searches for new consumers locally, meaning that the search originates from where the firm itself is located. Second, the firm uses its existing network of consumers to search remotely, meaning that the search originate from where the existing consumers are located. This second remote search captures the idea of local externalities as in the case of the geographic expansion of Wal-mart in Jia (2008) and Holmes (2011), or in the case of Chilean exporters in Morales, Sheu and Zahler (2011). It may either correspond to the technological constraint on the expansion of a distribution network as in Holmes (2011); to the cost of customizing a product for local tastes and requirements as in Morales, Sheu and Zahler (2011); or more generally to the notion that exporting entails some amount of traveling and communicating with business partners, so that a firm that exports to a location y will acquire some partial knowledge about that location y .

Before describing the dynamic acquisition of consumers, it is useful to introduce a few notations. A firm of age t located in the origin has a network of consumers in various locations. The total measure of consumers is m_t , and they are distributed over the real line \mathbb{R} according to the p.d.f. g_t . The function $f_t(x)$ then describes the density of consumers in location x ,

$$f_t : \mathbb{R} \rightarrow \mathbb{R}^+ \text{ with } \int_{\mathbb{R}} f_t(x) dx \equiv m_t \text{ and } g_t = \frac{f_t}{m_t} \quad (1)$$

so that the firm has a mass $\int_a^b f_t(x) dx$ of consumers in the interval $[a, b]$. The function f_t specifies both the number (density) and the geographic location of all the consumers of the firm. Note that f_t is not a probability density, as it sums up to m_t and not 1. The normalized g_t is a p.d.f.

The distribution of consumers f_t evolves as follows.

First, a firm searches locally for consumers from where it is located (the origin). The firm randomly draws a measure $\gamma\mu$ of consumers, where γ is the (constant) growth rate of the population of firms, and μ is a parameter.⁴ The geographic location $x \in \mathbb{R}$ of these consumers is random, and drawn from the symmetric p.d.f. $g(x)$. The distribution $g(x)$ is a measure of the geographical bias of this local search technology. Note that beyond symmetry, I do not impose any restriction on the specific shape of this distribution. While one may conjecture that contacts are more likely to be located close-by than far-away, so that the density $g(x)$ is expected go down as one moves away from the origin, I do not impose any such restriction. The density $g(x)$ may be non-monotonic, it may have mass points, or be zero over some range.

⁴Expressing the measure of randomly drawn new consumers as a multiple of the population growth rate is just a normalization, which will simplify the main results.

Second, given that a firm already has consumers in various locations, it searches for new consumers remotely from these locations. For each existing consumer in location $y \in \mathbb{R}$, the firm draws a measure $\gamma\mu\pi$ of consumers with $\pi \in [0, 1]$ a parameter. The geographic location of these consumers is drawn from the p.d.f. $g(x - y)$. Remote search works exactly as local search, except that (i) it is shifted from the origin to location y , and (ii) the efficiency of this remote search is discounted by a constant probability π .

Without loss of generality, neither does a firm lose consumers, nor do firms die. Adding a random death process to either contacts or firms does not change any of the results below, beyond some simple rescaling of the parameters.

The dynamic evolution of the network of consumers described above can be summarized in the following difference equation for f_t ,

$$f_{t+1}(x) - f_t(x) = \gamma\mu g(x) + \gamma\mu\pi \int_{\mathbb{R}} f_t(y) g(x - y) dy \quad (2)$$

with the initial condition $f(x) = 0, \forall x \in \mathbb{R}$. The change in the density of consumers in location x from time t to time $t + 1$ can be decomposed in two terms. The first term corresponds to the local search for new contacts. It brings a total of $\gamma\mu$ new contacts, with a density $\gamma\mu g(x)$ of them in x . The second term corresponds to the remote search for new contacts. Each of the existing $f_t(y) dy$ contact in location y offers the possibility for a remote search of $\gamma\mu\pi$ new contacts originating from y ; this remote search generates a total of $\gamma\mu\pi$ new contacts, with a density $\gamma\mu\pi g(x - y)$ of them in x ; since the remote search can be intermediated via any location y , the new consumers $\gamma\mu\pi f_t(y) g(x - y) dy$ found in x via y have to be integrated over all possible location $y \in \mathbb{R}$.

While the recursive definition of f_t in Equation (2) is complex, it always admits an analytical solution, and it admits a closed form solution for the special cases where g is either a Gaussian or a Cauchy distribution. In the interest of concision, I relegate these explicit solutions to Appendix A. These solutions however allow me to derive closed-form solutions for several quantities that can be measured in the data, thus offering a direct test of the model. In the next section, I analyze the distribution of the total number of consumers (m_t) within the population, as well as moments of the geographic distribution of these consumers.

1.2 The number and geography of consumers

The following proposition characterizes the invariant distribution of the total number of consumers.

Proposition 1 *For short time intervals, the fraction of firms with fewer than m consumers is,*

$$F(m) = 1 - \left(\frac{1}{1 + \pi m} \right)^{\frac{1}{\mu\pi}}$$

Proof. See Appendix A. ■

Let me briefly describe the properties of the cross-sectional distribution of the number of consumers, and provide some intuition.

First, note that geography plays no role in this distribution. This is due to the fact that the geographically biased distribution $g(\cdot)$ affects the location of consumers, but not the total number of them. Integrating Equation (2) over \mathbb{R} , I get a simple difference equation for the total number of a firm's consumers,

$$m_{t+1} - m_t = \gamma\mu + \gamma\mu\pi m_t \tag{3}$$

This process does not depend on any of the properties of the distribution $g(\cdot)$.

Second, the upper tail of the distribution asymptotes to a scale-free Pareto distribution, whereas the lower tail is close to an exponential distribution. This dimension of the model is very close to the model of acquisition of a network of “friends” in Jackson and Rogers (2007), which itself is an extension of the Steindl (1965) model of the firm size distribution.⁵

For m large, i.e. for firms that already have many consumers, the local search component ($\gamma\mu$) becomes negligible, and only the remote search component remains. Each existing consumer allows the firm to gain a constant number ($\gamma\mu\pi$) of new consumers. The growth rate of the number of consumers is approximately constant. This means that for m large, m grows approximately exponentially ($m_t \approx e^{\gamma\mu\pi t}$). Given that the population of firms grows exponentially at a rate γ , this leads to a Pareto distribution with an exponent $-1/(\mu\pi)$ in the upper tail.

For m small on the other hand, i.e. for firms with few consumers, the remote search component becomes negligible, and only the local search component remains. Each period, an approximately constant number ($\gamma\mu$) of new consumers are added. This means that for m small, m grows approximately arithmetically ($m_t \approx \gamma\mu t$). Given the exponential growth of the population of firms, this leads to an exponential distribution with parameter $1/\mu$ in the lower tail.

For intermediate values of m , the cross-sectional distribution of the number of consumers is a mixture of the above exponential and Pareto distributions. Plotting the counter-cumulative distribution $1 - F(m)$ in a log-log scale, the upper end would asymptote a straight line (the

⁵For more elaborate models of the dynamic evolution of size, see for instance Gabaix (1999), Luttmer (2007) or Rossi-Hansberg and Wright (2007).

Pareto upper tail), while the lower end would exhibit some degree of concavity (the exponential lower tail). Both the slope of the upper tail, the range over which the distribution is concave, and how concave it is, depend on the parameters μ and π . Observing the empirical distribution $F(m)$ allows to both test proposition 1 and estimate the parameters μ and π . This empirical test is presented in Section 2.2.

The recursive characterization of the distribution of consumers f_t in Equation (2) not only allows me to solve for the cross-sectional distribution of the total number of consumers, but also for any moment of the geographic distribution of these consumers. I define $\Delta(m)$ as the average (squared) distance from a firm's consumers for a firm with m consumers. I call f_m the density of consumers for a firm with m consumers, and $g_m = f_m/m$ its p.d.f. As the distribution g_m is symmetric around zero for any m , this average (squared) distance corresponds to the variance of the p.d.f. g_m ,

$$\Delta(m) \equiv \int_{\mathbb{R}} x^2 g_m(x) dx \quad (4)$$

The next proposition describes the relationship between the number of consumers of a firm, and the distance from these consumers.

Proposition 2 *For short time intervals, the average (squared) distance from a firm's consumers, $\Delta(m)$, increases with the number of consumers m ,*

$$\Delta(m) = \left(1 + \frac{1}{\pi m}\right) \ln(1 + \pi m) \Delta_g$$

with $\Delta_g \equiv \int_{\mathbb{R}} x^2 g(x) dx$ the variance of the p.d.f. $g(\cdot)$.

Proof. See Appendix A. ■

Over time, not only does a firm acquire more consumers, but the geographic distance from these consumers increases. This result is entirely due to the remote search, and can be understood as follows. Each time a firm gains one more consumer, it searches remotely from where this consumer is located. On average, existing consumers are some distance away from the firm, and remote searches bring new consumers that are themselves some distance away from the existing consumers. So each new wave of remote searches bring new consumers that tend to be further and further away.

Formally from the difference Equation (2), the location of the new consumers acquired via remote search is the sum of the signed distance of the existing consumers and the signed distance

of the remote search: for each consumer at a signed distance y , if the remote search delivers a new consumer that is herself at a signed distance $(x - y)$ from y , the new consumer will be at a signed distance $y + (x - y) = x$. In other words, the remote search process is equivalent to taking the sum of two random variables, the first one being the variable that describes the location of existing consumers, and the second one being the remote search. The fact that the variance of the sum of two random variables is the sum of their variances explains why the average (squared) distance from a firm's consumers increases over time. This can be seen formally in Equation (2). The term $\int_{\mathbb{R}} f_t(y) g(x - y) dy$ is the convolution product of the functions $f_t(\cdot)$ and $g(\cdot)$. In probability theory, the convolution product is used to study the sum of random variables: the p.d.f. of the sum of two random variables is the convolution product of their respective p.d.f.'s. This is the essence of the proof of Proposition 2.

From the reasoning above, and as can readily be seen in Proposition 2, the fact that the average (squared) distance from a firm's consumers increases with the number of consumers is only driven by the remote search process. Absent this remote search ($\pi = 0$) the average (squared) distance from consumers is constant: $\Delta(m) = \Delta_g, \forall m$. Without remote search, a firm accumulates over time more and more consumers from a series of independent waves of local searches. But all waves of new consumers brought by this local search have the exact same geographic distribution. Large firms sell more, but they have the same geographic distribution of sales as small firms. If remote search is present ($\pi > 0$), the average (squared) distance of sales increases with the number of consumers. Initially, for m small, the majority of new consumers come from local searches, and $\Delta(m)$ is relatively insensitive to m : $\partial \Delta(m) / \partial m|_{m=0} = 0$. As the number of consumers gets large, i.e. m large, the average (squared) distance of exports increases with the number of consumers in a log-linear way: $\Delta(m) \underset{m \rightarrow \infty}{\approx} \text{constant} + \Delta_g \ln(m)$.

Note that Proposition 2 holds for any arbitrary symmetric p.d.f. $g(\cdot)$ with a finite variance. This is true despite the fact that the geographic distribution of new consumers depends in a complex non linear fashion on the entire distribution $g(\cdot)$, and hence on all the moments of this p.d.f. This result, while striking at first, can easily be understood as follows. The average (squared) distance from a firm's consumers corresponds to the variance of the p.d.f. $g_t(\cdot)$. The variance of a distribution can be derived from the second derivative of the characteristic function of that distribution evaluated at zero. The proof of Proposition 2 in the Appendix shows how to use the difference Equation (2) to express the characteristic function of g_t in terms of the characteristic function of g . Then, for the same reason that the n^{th} derivative of a composition of functions

depends only on the first n derivatives of these functions, the variance of g_t (the 2^{nd} derivative of its characteristic function) does not depend on any moment of g higher than its variance (any derivative of its characteristic function of order above 2). The fact that the details of the geographic bias in the local search technology $g(\cdot)$ do not matter for Proposition 2 will prove extremely useful in guiding my empirical strategy. To test this proposition, I only need to estimate a single moment of the distribution $g(\cdot)$, but I do not require any additional information on the distribution of the geographic bias in local search $g(\cdot)$.

Using the same analytical tools, I can potentially describe all the moments of the distribution of consumers f_t . While I will not describe all these moments, I will make two observations that help understand the process of acquiring consumers.

First, the process of acquiring a network of consumers exhibits a strong history dependence. I have characterized above the underlying probability distributions that characterize the distribution of a firm's consumers. Any two firms in the same location and of the same age will have the exact same probability distribution of consumers at any point in time. But of course, this is only a continuous approximation of the discrete draws from such distributions. In a discrete version of the model, if a firm initially happens to gain consumers in one particular location, it is subsequently more likely to keep gaining consumers in the vicinity of that location. So over time, the distribution of consumers of two initially identical firms will tend to diverge, each following its own history dependent path. In Equation (2), the location of a firm's new consumers at time $t + 1$ depends on the location of its existing consumers at time t . I use this observation to test empirically Equation (2).

Second, over time, not only is a firm able to reach more consumers that are further away, but the geographic dispersion of these contacts increases as well. This result is also due to the history dependence of the search process, but this time within and not between firms. Each existing consumer allows a firm to acquire new consumers that will tend to be geographically concentrated around these existing consumers. Consumers tend to be clustered around each other. Time brings new clusters of consumers that tend to be increasingly far apart from each other. I derive this result formally and test this prediction in the online appendix.⁶

In this section, I have shown how to formally analyze a process of firms acquiring a network of consumers that features two properties: first, the search for consumers tends to be geographically

⁶See http://home.uchicago.edu/tchaney/research/TradeNetworks_Appendix.pdf

biased, and second, existing consumers allow a firm to remotely search for new consumers in the vicinity of the existing ones. In the following section, I show that my results are robust to relaxing the strong assumptions that warranted the tractability of the analysis above, and I discuss the relationship of my model with the existing literature.

1.3 Robustness and discussion

Robustness: I make two strong simplifying assumptions in the model. First, I assume that the world is infinite (represented by the real line \mathbb{R}) and uniformly populated. Obviously, this assumption is at odds with the data. The world is finite. Actual countries in that world are not uniformly distributed: they vary in size and population density, and the world contains many empty areas (oceans). Second, I assume that the geographically biased search function $g(\cdot)$ is the same in any location. No matter where (for any $y \in \mathbb{R}$) a firm acquires consumers, the remote search from these locations is the same ($g_y(x - y) = g_z(x - z) = g(x)$ for any $(x, y, z) \in \mathbb{R}^3$). Obviously, this assumption is again at odds with the data. Some countries are isolated (remote islands), so that no consumer can be found in their vicinity. Some countries are in the center of densely populated areas (continents). And some countries are at the edges of such populated areas, so that consumers can only be found in one direction and not the other.

While these simplifying assumptions are required to keep the model tractable and to derive analytically the main propositions, I show numerically that they are inconsequential. In other words, I show that the model under the uniformity and symmetry assumptions is a precise approximation of a more complicated model (which I can only solve numerically) without these assumptions.

To do so, I randomly generate a discrete non-uniform population distribution over a finite interval.⁷ The artificial world I simulate features areas of various sizes (islands and continents), separated by empty spaces of various sizes (mountains, deserts and oceans), as well as a varying population density (large and small cities). Following the literature on city sizes as in Gabaix (1999), I assume that the population of firms in each city grows at the same constant rate γ . This is a special case of the home market effect, whereby new firms tend to cluster in larger cities.

After a firm is born, it randomly meets consumers in other cities. I depart from the symmetry assumption of the $g(\cdot)$ function, and assume instead that the search process originating from each city is different (generically, $g_y(x - y) \neq g_z(x - z)$ for $y \neq z$). Not only are the various $g_y(\cdot)$

⁷Detailed instructions for replicating this numerical simulation are given in the online appendix.

See http://home.uchicago.edu/tchaney/research/TradeNetworks_Appendix.pdf

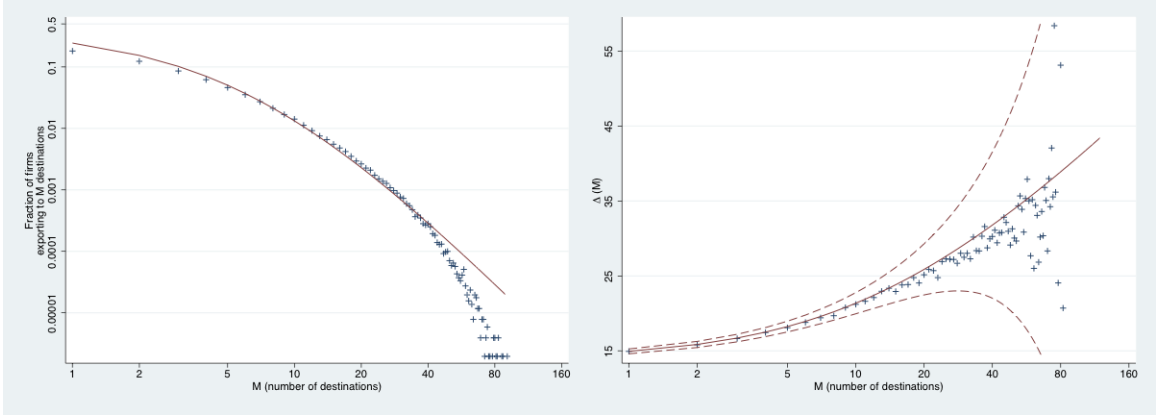


Figure 1: Numerical simulations of Propositions 1 and 2.

Notes: The solid red line in the left panel corresponds to the theoretical prediction from Propositions 1, and in the right panel to the theoretical prediction from Proposition 2. The dotted lines on the right panel correspond to the predicted 90% interval band. The blue plus signs correspond to the results from the numerical simulation.

functions different, but even their variances Δ_{g_y} differ. They are the same on average, but there is some dispersion around this average. Finally, I consider a discrete version of the model above. In a sense, I also test how far the law of large numbers is from applying in this discrete model.

The results of these numerical simulations for Proposition 1 and 2 are presented in Figure 1. The numerical simulations show that the convenient assumptions that the world is unbounded, continuous and uniformly populated, and that all locations face the same distribution g are not crucial for the main theoretical predictions of the model. On the left panel, the simulated cross-sectional distribution of the number of consumers is close to the theoretical prediction in Proposition 1. On the right panel, the average (squared) distance from a firm's consumers closely matches the theoretical prediction of Proposition 2. Among firms with many consumers, the simulated data exhibits increasing noise. This is due to two forces. First, there are few firms with many consumers, so that the average realization among these small samples has a high variance. In addition, as discussed above, the variance of the (squared) distance from any firm's consumers increases with the number of consumers. Among firms with many consumers, the small sample accounts for a quarter of the noise, and the increased variance of (squared) distances for each firm the remaining three quarters. This pattern of the simulated data resembles the actual data on firm level exports presented in Section 2.3.

Relation to the existing firm-level trade literature: Existing international trade models with heterogeneous firms, such as Bernard, Eaton, Jensen and Kortum (2003) or Melitz (2003)

and its extension in Chaney (2008) do not offer specific predictions regarding the distribution of the number of countries reached by different firms. By comparison, the model I develop offers a parsimonious theory for the extensive margin of international trade.

For this discussion, I assume that in my model, the number of contacts of a firm, m , is proportional to the number of countries it exports to, M : $m \propto M$. In the empirical analysis in the next section, I discuss this approximation in details and offer a statistical correction for it.

In the original Melitz (2003) model, all trade barriers are symmetric, and any exporter exports to all foreign markets. This is obviously an artifact of the simplifying yet counter-factual assumption that all trade barriers and country sizes are perfectly symmetric. In Chaney (2008), I offer a simple extension of Melitz (2003) with asymmetric country sizes and fixed and variable trade barriers. In this model, from the point of view of a given exporting country, say France, there is a strict hierarchy of foreign markets. This means that markets can be strictly ordered in a decreasing level of accessibility, so that if a French firm exports to the M^{th} most accessible market, it will necessarily export to all markets $M' \leq M$. Therefore, the fraction of firms that export to exactly M markets is simply the fraction of firms that have a productivity between the productivity threshold for exporting to market M and the threshold for exporting to market $M + 1$. Even if productivities are distributed Pareto, the fraction of firms that export to exactly M markets can take any value, depending on the distance between the thresholds for exporting to country M and $M + 1$. Even if country sizes are themselves Pareto distributed, and if fixed export costs are log-proportional to country size, there is no reason to make the counterfactual assumption that variable trade barriers are themselves log-proportional to country size. The fraction of firms that export to exactly M markets does not even have to be decreasing in M .⁸

By adding to the Melitz/Chaney model firm-destination specific idiosyncratic shocks to the entry cost and demand faced by each firm, Eaton, Kortum and Kramarz (2011) can a priori replicate any pattern of entry in the data. Calibrating their model to the data, they need to assume a very large amount of idiosyncratic noise, with a ratio of the relevant combination of fixed entry cost and local demand shocks of 1 to 13 between the 25th and the 75th percentiles. So the productivity thresholds are essentially randomly distributed. With the assumption of this additional noise, the fraction of firms that export to exactly M markets inherits the assumed Pareto distribution of productivities across firms, which matches the data well. This distribution

⁸I develop these arguments formally and provide a calibration of the Melitz/Chaney model on the same data I use in Section 2 in an online appendix.

See http://home.uchicago.edu/tchaney/research/TradeNetworks_Appendix.pdf.

is directly assumed, and the fact that the model lines up with the data comes from the assumption of a large amount of idiosyncratic noise and of Pareto distributed productivity shocks, and not from the underlying Melitz/Chaney model.⁹ In contrast, the network model I develop offers a theory of the distribution of entry into foreign markets, without the need for ad hoc assumptions on the firms' productivity distribution.

In the stochastic model of Bernard, Eaton, Jensen and Kortum (2003), there is no strict hierarchy in the accessibility of foreign markets. A given exporter, even if it has a low productivity, may still export to many foreign countries, if this exporter is lucky enough to face unproductive foreign competitors. However, the structure of country sizes, relative productivities and labor costs across countries, and bilateral trade barriers between countries imposes a severe restriction on the cross-sectional distribution of the number of foreign markets entered. For a large number of firms, or for the continuous limit developed in the model, there is no uncertainty either in the fraction of firms entering any given market, or the distribution of the number of markets entered. This distribution depends on the specific trade barriers and country characteristics. Even with the assumed ad hoc and convenient Fréchet distribution of productivities, there is no reason why any particular distribution should arise. As in the Melitz model, the fraction of firms that export to exactly M markets does not even have to be decreasing in M . The following argument makes this point clear. In the limit of infinitely large trade barriers, all firms only sell in their domestic market, so that no firm sells to any $M > 0$ foreign markets. In the other extreme of perfectly free trade, all firms that sell domestically also export to all countries in the world. So whereas the fraction of firms that export to all foreign countries in the world is monotonically decreasing from 1 to 0 with the level of trade barriers, the fraction of firms exporting to any other number of foreign countries is not monotone. The fraction of firms exporting to exactly M markets can be made arbitrarily small or large by simply varying bilateral trade barriers.

Finally, if trade barriers increase with distance, and if there is no systematic correlation between country size and distance from France, both the Melitz/Chaney model and Bernard, Eaton, Jensen and Kortum (2003) would correctly predict that the distance of exports increases with the number of markets a firm enters. However, neither model offers any specific prediction for the shape of this relationship. Even if a large amount of noise is added as in Eaton, Kortum and Kramarz (2011),

⁹Similarly, Armenter and Koren (2010) estimate from the data the distribution of the number of shipments (the distribution of the number of "balls") from the data, and then generate predictions for the occurrence of zeroes in the trade data (empty "bins"). By contrast, instead of assuming this distribution to match the data, my model offers a theory that generates such a distribution.

the very strong tendency of firms in the Melitz/Chaney model to first enter close by markets implies that exports are far more geographically concentrated than in the data. For instance, among firms that export to a single foreign market, the average squared distance (in thousands of km) between France and that country is 18 in the data, 16 in my calibrated model, but only 2 in the calibrated Eaton, Kortum and Kramarz (2011) model.¹⁰¹¹

To summarize, while existing firm level trade models directly make ad hoc assumptions to match the extensive margin of trade, I develop a parsimonious model that endogenizes these assumptions. On the other hand, my model is silent about the determinants of the intensive margin of trade, or about the relation between a firm's exposure to international trade and its size in different markets, while those models make precise predictions about those. In that sense, the proposed network model is complementary to the existing firm level trade literature.

2 Empirical evidence

In this section, I bring several the key testable predictions from the theoretical model to the data. In Section 2.1, I describe the data on firm level exports for French firms, as well as aggregate bilateral trade flows for the rest of the world. In Section 2.2, I test the first main prediction of the model regarding the cross-sectional distribution of entry into different foreign markets, derived from Proposition 1. In Section 2.3, I test the second main prediction of the model regarding the geographic dispersion of exports across firms, derived from Proposition 2. In Section 2.4, I test directly the assumption of the model regarding the dynamic entry of firms into foreign market, derived from Equation (2). In doing so, I link formally the time-series and the cross-section of firm level exports.

2.1 Data

To bring the model to the data, I use two sources of data. First, I use firm level export data for French exporters, over the period 1986-1992. The data used come from the same source as the data used by Eaton, Kortum and Kramarz (2011). I use information on French exporters in

¹⁰For a more intuitive interpretation of these numbers, the average distance is 3,500 km in the data versus 900 km in the calibrated Eaton, Kortum and Kramarz (2011) model.

¹¹Of course, as in the Melitz/Chaney model, as firms eventually enter all countries, the difference between the data and the model shrinks. Among firms that export to the maximum observed number of countries (98), the average squared distance is 38 in the data, 43 in my calibrated model, versus 36 in the Eaton, Kortum and Kramarz (2011) model. Those numbers however are not precisely estimated, as few firms export to that many markets.

the years 1986 to 1992. For each firm, I know the total value (in French Francs) of its exports over a given year, to a given country. There are between 119,000 exporters (in 1988) and 130,000 exporters (in 1987) in my sample (126,594 in 1992). Those firms export to a total of 210 different foreign countries. French exporters export on average to between 3.8 (in 1991) and 4.2 (in 1986) different foreign markets (3.9 in 1992).

In addition to these data on firm level exports for France, I use information on the size of countries, their distance from France and from each other, and aggregate bilateral trade between country pairs. The size of a country is measured as nominal GDP, collected from the Penn World Tables. The distance between two countries is the population weighted geodesic distances between the main cities in both countries, which come from the CEPII. Finally, I use data on aggregate bilateral trade flows between countries, which are collected from the NBER.

2.2 Matching the distribution of export destinations

In this section, I test the first main prediction of the model, Proposition 1. The model predicts that the distribution of the number of consumers of a given firm can be described by a mixture of an exponential and a Pareto distribution. The only two parameters that govern this distribution are μ , the number of new consumers acquired each period via local search (expressed in multiples of the firm population growth rate γ), and π , the efficiency of remote search relative to local search. As a reminder, the theory predicts that the fraction of firms that have fewer than m consumers is,

$$F(m) = 1 - \left(\frac{1}{1 + \pi m} \right)^{\frac{1}{\mu\pi}}$$

There is one main complication that arises when bringing this prediction to the data: the firm level export data only provides information on the number of foreign countries a firm exports to, not the number of consumers it sells to. While the underlying number of contacts, m , is not observed, the observed distribution of the number of countries, M , provides enough information to recover the parameters that govern the distribution $F(m)$. I follow the guidance of the theory and ask: given that the number of contacts, m , is distributed according to the distribution $F(m)$, and that there are N equal sized countries, what is the distribution of the number of countries, M , accessed by different firms? The following proposition answers this question.

Proposition 3 *If there are N equal size distinct foreign countries, and the distribution of the number of consumers across firms is given by Proposition 1, then the fraction of firms that export*

to exactly M distinct countries is,

$$\phi(M|N) = \sum_{m=M}^{+\infty} \frac{1}{\mu} \left(\frac{1}{1 + \pi m} \right)^{\frac{1}{\mu\pi} + 1} n! S_2(m, M) \binom{N}{M} \left(\frac{1}{N} \right)^m$$

with $S_2(m, M)$ the Stirling number of the second kind.

Proof. See Appendix A. ■

The distribution $\phi(M|N)$ of the number M of countries firms export to depends on the same parameters μ and π as the distribution $F(m)$ of the number m of contacts, as well as the known total number of countries, $N = 210$. It is essentially the same distribution, filtered through the coarse grid of the 210 countries of the world. The data on firm level exports gives me as many realizations of the distribution $\phi(M|N = 210)$ as there are firms in the sample (126,594). Using a maximum likelihood procedure, I can estimate the parameters of interest, μ and π .

I make two strong simplifying assumptions. First, I assume that all countries have the same size. Second, I assume that geography does not distort the mapping between the distribution of the number of contacts, $F(m)$, and the number of distinct countries, $\phi(M|N = 210)$, firms export to. Note that geography (summarized in the p.d.f. g) has no impact on the distribution of the total number of contacts, as the number of contacts evolves recursively according to Equation (3) irrespective of geography. However, geography does affect the likelihood of a contact falling into another distinct country. While I could in theory use the actual sizes and geographic locations of all countries worldwide and the guidance of the model, incorporating this vast amount of information into the estimation is numerically infeasible. Instead I simulate a discrete version of the model in a world with asymmetric country sizes, and a complex geographic structure and show that ignoring asymmetries in country sizes and geography is inconsequential: when I perform the maximum likelihood estimation procedure described above on the simulated data, I recover parameters that are close to the true parameters used to generate the data.¹²

The results are presented in column (1) of Table 1. The empirical cross sectional distribution of entry into different foreign markets by French exporters suggest that $\mu = 2.38$ and $\pi = 0.158$. In other words, remote search is about 6 times less efficient than direct search ($\pi = 1/6.32$). Of course, as firms acquire more contacts, remote search accounts for an increasing share of the firm's new contacts. For a firm with the sample mean number of 3.9 foreign contacts, direct search is only

¹²The details of this numerical simulation and the results from the estimation are given in the online Appendix. See http://home.uchicago.edu/tchaney/research/TradeNetworks_Appendix.pdf.

Table 1: Empirical fit of Proposition 3

	(1) (MLE w. correction)	(2) (MLE)	(3) (NLLS)	(4) (MLE)
π	0.158 (.0017)	0.148 (.0017)	0.130 (.021)	0
μ	2.384 (.0098)	2.394 (.0098)	2.970 (.349)	3.84 (.011)
Adj. R^2/\log (likelihood.)	-282,098	-282022	0.976	-296,916
likelihood ratio test: (4) vs. (2)	$\Lambda = 29,788$, p-value < .0001			

Notes: This table presents the estimates of parameters μ and π using different procedures. These two parameters govern the distribution of the number of foreign contacts in Proposition 1 and 3. Column (1) uses Proposition 3 to correct for the fact that the number of countries, M , and not of contacts, m , is observed. Columns (2)-(4) use Proposition 1 under the simplifying assumption $M = m$. I use the same data on French exporters in 1992 for all procedures. Standard errors are in parentheses. All coefficients are statistically different from zero at the 1% level of significance.

60% more important than remote search. For a firm with 20 foreign contacts (90th percentile), remote search dominates, and accounts for more than 75% of new contacts. On average, remote search accounts for 21% of all new contacts.¹³ Since firms with many contacts are also very large exporters in value, from a macro-economic perspective, the formation of contacts is dominated by remote search. Figure 2 plots the empirical density of the number of foreign markets served by French exporters and the theoretical prediction from Proposition 3 with the estimated parameters μ and π . The fact that both random and network-based meetings coexist explains the curvature of the empirical density in a log-log scale. The flattening out of the curve for very large observations is due to the fact that once a firm already has contacts in most countries in the world, it becomes less and less likely that new contacts fall into one of the few yet unaccessed countries.

For robustness, I also estimate the parameters (μ, π) through various alternative specifications. The parameter estimates vary little across the various specifications or estimation methods. In column (2) of Table 1, I estimate the parameters through Maximum Likelihood under the simplifying assumption that a firm that exports to exactly M distinct foreign countries has exactly M differ-

¹³I use a continuous approximation to circumvent integer constraints. In any period, direct search delivers $\bar{m}_{direct} = \gamma\mu$ new contacts for each firm, irrespective of their number of existing contacts. Remote search brings $\gamma\mu\pi m$ new contacts for a firm with m existing contacts. Given that the contacts are distributed within the population according to the c.d.f. $F(m)$, the average number of new contacts brought by remote searches is $\bar{m}_{remote} = \int_0^{+\infty} \gamma\mu\pi m dF(m) = \gamma\mu(\mu\pi)/(1 + \mu\pi)$. The fraction of new contacts delivered by remote search is therefore $\bar{m}_{remote}/(\bar{m}_{direct} + \bar{m}_{remote}) = \mu\pi/(1 + 2\mu\pi) \approx .21$.

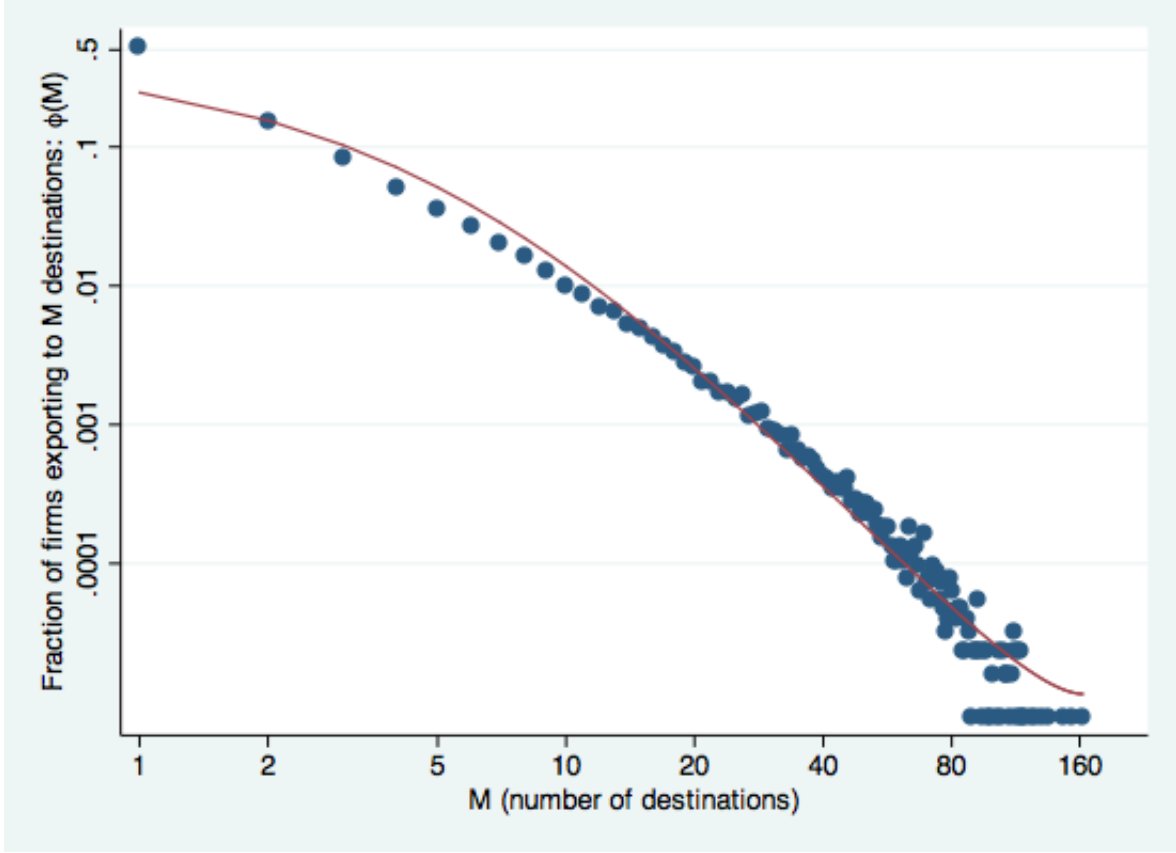


Figure 2: Empirical fit of Proposition 3, $\phi(M)$ versus M .

Notes: $\phi(M)$ is the fraction of firms that export to M distinct destinations; dots: data, all French exporters in 1992; line: theory. $\pi = .158$ (.022) and $\mu = 2.38$ (.019) are estimated through Maximum Likelihood.

ent foreign contacts. In column (3), I estimate through a Non Linear Least Square procedure the parameters under the same simplifying assumption that one country corresponds to one contact.¹⁴ Finally, in column (4), I estimate through Maximum Likelihood a simpler benchmark model where contacts are only acquired from direct search, again under the same simplifying assumption that one country corresponds to one contact. I propose a likelihood-ratio test of my model against this simpler benchmark, and reject the simpler model unambiguously. I offer further robustness checks in the Appendix.¹⁵

The fact that the estimated parameters vary little across the different specifications is not

¹⁴Specifically, I estimate through NLLS the parameters (α, μ, π) in the following equation,

$$\ln(\text{fraction of firms exporting to } M \text{ countries}) = \alpha - \left(1 + \frac{1}{\mu\pi}\right) \ln(1 + \pi M) + \epsilon$$

where ϵ is a normally distributed error term. This corresponds to the log of Proposition 1.

¹⁵See Table 5 in Appendix B.

surprising. Eaton, Eslava, Krizan, Kugler and Tybout (2010) are able to identify separately each U.S. contact of Colombian exporter. They show that 80% of Colombian firms that export to the U.S. have a single contact (buyer) there. In the context of within country trade, using detailed information on the input-output linkages between individual US firms, Atalay, Hortaçsu, Roberts and Syverson (2010) report that the average number of U.S. suppliers of U.S. firms is only marginally above 1.

Armed with an estimate for the relative importance of direct versus remote search, π , I study the geographic dispersion of exports across firms in the next section.

2.3 Matching the geographic dispersion of exports

In this section, I test the second main prediction of the model, Proposition 2. The model predicts the specific way in which the geographic dispersion of exports increases as firms enter more foreign markets. This relationship only depends on the relative efficiency of remote and direct search, π .

Using data on the geographic distribution of exports among firms exporting to exactly M foreign markets, I construct an empirical measure of the geographic dispersion of exports, $\Delta(M)$. To construct $\Delta(M)$, I need to measure both the geographic distance between France and other countries, as well as the p.d.f. g_M which governs the entry of firms into different countries. For all firms with M contacts, exports are distributed according to the p.d.f. g_M . The probability that a firm exports to country c is given by the integral of p.d.f. g_M over all locations $x \in [\underline{c}, \bar{c}]$ that belong to country c ,

$$\Pr[\text{exports}_{M,c} > 0] = \int_{\underline{c}}^{\bar{c}} g_M(x) dx$$

To derive an empirical measure of g_M , I use two observations. First, the probability of exporting to a country is approximately proportional to the measure of locations inside c , $(\bar{c} - \underline{c})$, which I approximate by the *GDP* of country c .¹⁶ Second, given many observations, I use the law of large numbers and approximate the probability that a firm with M contacts exports to country c by the fraction of firms that do so. Calling c the center of country c , I get the following empirical measure of $g_M(c)$,

$$\widehat{g_M(c)} \propto \sum_{i \in E(M)} \left(\frac{1}{GDP_c} \right) \mathbb{I}\{\text{exports}_{i,c} > 0\} \quad (5)$$

¹⁶The results below are not sensitive to removing the control for *GDP*, and are stable over time, as shown in the robustness checks in Table 6 in Appendix B.

where $E(M)$ is the set of firms that export to M distinct countries. I can now go back to the definition of the average (squared) distance of exports among firms with M contacts. I use the approximation that within country distances are small relative to between country distances for the last equality, and get,

$$\Delta(M) \equiv \int_{\mathbb{R}} x^2 g_M(x) dx = \sum_c \int_{\underline{c}}^{\bar{c}} x^2 g_M(x) dx \approx \sum_c (Distance_{Fr,c})^2 g_M(c) \quad (6)$$

Combining Equations (5) and (6) above, I define the empirical counterpart to $\Delta(M)$ as follows,

$$\widehat{\Delta(M)} = \frac{\sum_{i \in E(M),c} (Distance_{Fr,c})^2 \left(\frac{1}{GDP_c} \right) \mathbb{I}\{export_{i,c} > 0\}}{\sum_{i \in E(M),c} \left(\frac{1}{GDP_c} \right) \mathbb{I}\{export_{i,c} > 0\}} \quad (7)$$

I can now bring Proposition 2 to the data. As a reminder, under the simplifying assumption that the number of countries a firm exports to proxies for its number of consumers, $m \approx M$, the theory predicts the following relationship between the geographic dispersion of exports and the number of markets a firm is able to enter,

$$\Delta(M) = \left(1 + \frac{1}{\pi M} \right) \ln(1 + \pi M) \times \Delta_g$$

Using the cross sectional distribution of the number of export destinations across firms, I estimated in the previous section that $\pi \approx .158$. I only need to calibrate Δ_g , which is not a unit-free measure, to bring the theoretical prediction to the data. I use a non linear least square estimation of the previous equation, and recover $\Delta_g \approx 14.860$ (.109), with an R^2 of 87%.¹⁷

Figure 3 plots the geographic dispersion of exports, $\Delta(M)$, as a function of the total number of foreign countries entered, both in the data and in the theory. Note that I only calibrate the intercept of this relationship. I have no other degrees of freedom to calibrate the shape of this relationship, which is entirely governed by the parameter π , estimated in the previous section on the cross-sectional distribution of the number of foreign destinations.

The theory connects two distinct empirical observations: the distribution of the number of foreign markets firms export to, and the geography of those exports. The same process for the formation of a network of foreign contacts governs both. The evidence presented in this and the previous section connects these two observations.

The next section directly tests some of the underlying assumptions of the model regarding the dynamic process of network formation.

¹⁷Each observation is weighted by the precision of its estimation. This precision is given by the square root of the number of observations used to estimate each second moment.

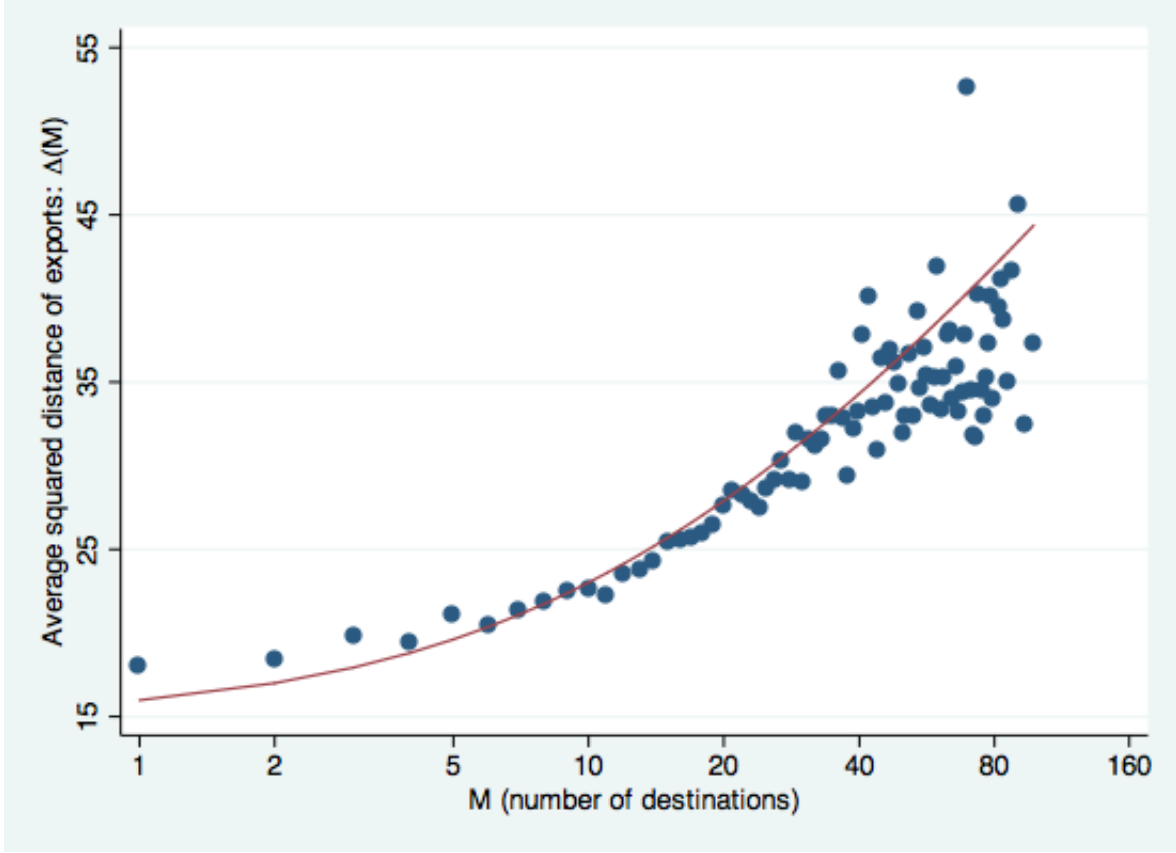


Figure 3: Empirical fit of Proposition 2, $\Delta(M)$ versus M .

Notes: $\Delta(M)$ is the average squared distance to a firm's export destinations, among firms exporting to M destinations, as defined in Equation (7); distances are calculated in 1,000's of kms; dots: data, all French exporters in 1992; line: theory. $\pi = .158$ is taken from the estimation of Proposition 3, and $\Delta_g = 14.860$ (.109) is estimated through non linear least squares, each point weighted by the square root of the number of observations used to compute $\Delta(M)$.

2.4 Matching trade dynamics

In this section, I directly test the main predictions of the theoretical model regarding the time-series of entry of individual firms into foreign markets. Firms use their existing foreign contacts to remotely search for new contacts. This implies not only that the more foreign contacts a firm already has, the more likely it is to enter new markets, but also that a firm follows a history dependent path when expanding into foreign markets: if firm i exports to country c' at time t , it is subsequently more likely to enter any country c that is closely connected to country c' .

To test those predictions, I formally bring Equation (2) to the data. The model presented in Section 1 is a model of trade between different cities (represented by different coordinates on the

real line), not explicitly a model of trade between countries. To bring the model to the data, I assume that countries form a partition of all locations in the world, with coordinates $x \in [\underline{c}, \bar{c}]$ belonging to country c . I set France as the origin country, and assign by convention all locations inside France to the origin ($x = 0$). Integrating Equation (2) over all locations in the interval $[\underline{c}, \bar{c}]$, i.e. in country c , I get the following equation for the dynamic of entry of an individual firm i into country c over time,

$$\int_{\underline{c}}^{\bar{c}} f_{i,t+1}(x) dx = \gamma\mu \int_{\underline{c}}^{\bar{c}} g(x) dx + \gamma\mu\pi \sum_{c'} \int_{y=\underline{c}'}^{\bar{c}'} f_{i,t}(y) \left(\int_{x=\underline{c}}^{\bar{c}} g(x-y) dx \right) dy + \int_{\underline{c}}^{\bar{c}} f_{i,t}(x) dx \quad (8)$$

While I do not have direct information about exports to a particular location within a country, I know whether a given firm exports or not to country c . I use this binary information as follows, with $\mathbb{I}\{export_{i,c,t} > 0\}$ an indicator function equal to 1 if firm i exports to country c in year t ,

$$\int_{\underline{c}}^{\bar{c}} f_{i,t}(x) dx = \mathbb{I}\{export_{i,c,t} > 0\} \quad (9)$$

Finally, I do not know a priori the exact shape of the geographically biased friction in the search technology, summarized by the function $g(\cdot)$. I use two alternative proxies for this search friction. With the first proxy, I follow a strict interpretation of the model and simply assume that this search friction is proxied by (the log of) geographic distance. This is a strict interpretation of the model in the sense that it assumes that the search friction is time invariant, symmetric, and the same irrespective of the origin of the search. Formally, if location y is in country c' and all locations $x \in [\underline{c}, \bar{c}]$ are in country c , I assume,¹⁸

$$\int_{\underline{c}}^{\bar{c}} g(x-y) dx \propto \ln Dist_{c',c} \quad (10)$$

With the second proxy, I follow a more flexible interpretation of the model and assume that the search friction is proxied by the actual aggregate trade flows between the two countries. Since distance is already a very good predictor of the level of trade, I use variations in trade flows as a proxy for the contemporaneous search friction.¹⁹ Formally, if location y is in country c' and location x is in country c , I assume,

$$\int_{\underline{c}}^{\bar{c}} g_t(x-y) dy \propto \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t}} \quad (11)$$

¹⁸Note that as for the measure $\Delta(M)$ above, I also add a control for GDP_c in the main empirical specification.

¹⁹Controlling directly for the level of aggregate bilateral trade flows does not affect substantially the results compared to simply controlling for distance.

Combining Equations (8), (9), (10) and (11), I estimate a Probit regression of different specifications of the following equation,

$$\begin{aligned}
\mathbb{I}\{export_{i,c,t+1} > 0\} = & \alpha \sum_{c'} \mathbb{I}\{export_{i,c',t} > 0\} \\
& + \beta_1 \ln Dist_{Fr,c} + \beta_2 \sum_{c'} \mathbb{I}\{export_{i,c',t} > 0\} \ln Dist_{c',c} + \beta_3 \sum_{c' \neq Fr} \ln Dist_{c',c} \\
& + \gamma_1 \sum_{c'} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t}} + \gamma_2 \sum_{c'} \mathbb{I}\{export_{i,c',t} > 0\} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t}} \\
& + \delta \mathbb{I}\{export_{i,c,t} > 0\} + Controls_{c,t} + \epsilon_{i,c,t} \quad (12)
\end{aligned}$$

The coefficient α controls for the simplest version of the model, where the probability of a firm entering a new country c is simply increasing in the number of countries this firm already exports to. The coefficients β_1 and γ_1 control for the cost of direct search. The coefficients β_2 and γ_2 control for the cost of remote search, which originates only from the countries where a firm is already exporting. The search cost is proxied by geographic distance for the coefficients β 's as in Equation (10), and by aggregate trade volumes for the coefficients γ 's as in Equation (11). Finally, the coefficient δ controls for the export status of firm i in the previous year, and the possibility that a firm loses foreign contacts. I expect $\alpha, \gamma_1, \gamma_2 > 0$, $\beta_1, \beta_2 < 0$ and $0 < \delta < 1$.

In addition, I control for country size, since it is mechanically more likely that a firm exports to a large country than to a small one. I also add controls for the sector in which a firm operates, as firms in different sectors may be more or less likely to export to any particular country. Removing the sector fixed effects does not affect the results. Finally, it is likely that if country c is more isolated from the rest of the world, in the sense that it is more distant from all other countries, competition in c will be relatively mild, and everything else equal, it will be easier to access c . The coefficient β_3 , expected to be positive, controls for this remoteness measure.

Table 2 shows the results of the Probit estimation of different specifications of Equation (12), and Table 3 shows the marginal effects of these regressions. In every specification, all coefficients are statistically significant (at the 1% confidence level), and of the expected signs.

More interestingly, the estimates from this panel regression are quantitatively close to the predictions of the model calibrated on the cross-sectional distribution of the number of contacts only. Using the marginal effect in column (1) of Table 3, the estimated increment in the probability of exporting to a given country due to adding an extra contact is .48%. Across the different specifications, this increment varies between .2% and .8%. Using the estimate for π from the

Table 2: Existing number of contacts and trade between third countries predict entry (PROBIT)

	Dependent Variable: $\mathbb{I}\{\text{export}_{i,c,t+1} > 0\}$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\sum_c \mathbb{I}\{\text{export}_{i,c',t} > 0\}$	0.0425 (.00024)	0.0545 (0.00035)	0.0318 (0.00025)	0.1562 (0.00180)	0.0323 (0.00026)	0.1659 (0.00184)
$\ln \text{Dist}_{F,c}$		-0.6053 (0.00167)	-0.3799 (0.00117)	-0.4429 (0.00239)	-0.3806 (0.00119)	-0.4445 (0.00219)
$\sum_{c'} \mathbb{I}\{\text{export}_{i,c',t} > 0\} \ln \text{Dist}_{c',c}$				-0.0143 (0.00020)		-0.0154 (0.00021)
$\sum_{c' \neq F} \ln \text{Dist}_{c',c}$				0.0057 (0.00007)		0.00606 (0.00007)
$\sum_{c'} \frac{\Delta \text{Exports}_{c',c,t}}{\text{Exports}_{c',c,t}}$					0.0599 (0.00233)	0.0565 (0.00229)
$\sum_{c'} \mathbb{I}\{\text{export}_{i,c',t} > 0\} \frac{\Delta \text{Exports}_{c',c,t}}{\text{Exports}_{c',c,t}}$					0.0159 (0.00155)	0.0067 (0.00150)
$\ln \text{GDP}_{c,t}$		0.2333 (0.00151)	0.1484 (0.00089)	0.1222 (0.00075)	0.1490 (0.00094)	0.1267 (0.00075)
$\mathbb{I}\{\text{export}_{i,c,t} > 0\}$			2.0740 (0.00364)	1.9991 (0.00353)	2.0528 (0.00363)	1.9700 (0.00355)
Constant	-2.1228 (0.00219)	-0.0951 (0.04955)	-1.2028 (0.05180)	-5.5761 (0.07119)	-1.2844 (0.05248)	-5.9416 (0.07235)
Sector Fixed Effect	yes	yes	yes	yes	yes	yes
N. obs		19,116,594			18,242,809	
N. firms		40,395			40,330	
N. years		6			6	
N. destinations		103			103	
Pseudo- R^2	0.1293	0.3567	0.5590	0.5572	0.5611	0.5597

Notes: This table shows the results of the PROBIT estimation of Equation (12) for a panel of all French exporters between 1986 and 1992. The dependent variable is an indicator function that takes the value 1 if firm i is exporting to country c at time t . The description of the explanatory variables is given along with Equation (12) on page 27. Standards errors are clustered at the firm level. All coefficients are statistically different from zero at the 1% level of significance.

Table 3: Existing number of contacts and trade between third countries predict entry (PROBIT: marginal effects)

Dep. Var.: $\mathbb{I}\{export_{i,c,t+1} > 0\}$	(1)	(2)	(3)	(4)	(5)	(6)
$\sum_{c'} \mathbb{I}\{export_{i,c',t} > 0\}$	0.0048 (0.00003)	0.0030 (0.00002)	0.0015 (0.00001)	0.0074 (0.00009)	0.0015 (0.00001)	0.0080 (0.00010)
$\ln Dist_{Fr,c}$		-0.0336 (0.00015)	-0.0176 (0.00007)	-0.0210 (0.00009)	-0.0180 (0.00007)	-0.0214 (0.00009)
$\sum_{c'} \mathbb{I}\{export_{i,c',t} > 0\} \ln Dist_{c',c}$				-0.0007 (0.00001)		-0.0007 (0.00001)
$\sum_{c' \neq Fr} \ln Dist_{c',c}$				0.0003 (0.00000)		0.0003 (0.00000)
$\sum_{c'} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t}}$					0.0028 (0.00011)	0.0027 (0.00011)
$\sum_{c'} \mathbb{I}\{export_{i,c',t} > 0\} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t}}$					0.0008 (0.00007)	0.0003 (0.00007)
$\ln GDP_{c,t}$		0.0129 (0.00007)	0.0067 (0.00004)	0.0058 (0.00003)	0.0070 (0.00004)	0.0061 (0.00003)
$\mathbb{I}\{export_{i,c,t} > 0\}$			0.4280 (0.00140)	0.4039 (0.00131)	0.4223 (0.00135)	0.3952 (0.00133)

Notes: This table shows the marginal effects for the PROBIT estimation of Equation (12) presented in Table 2. The marginal effect is calculated as dy/dx at the average value of each x in the sample. dy/dx is for a discrete change from 0 to 1 when x is a dummy variable. Standards errors are clustered at the firm level. All coefficients are statistically different from zero at the 1% level of significance.

estimation of the cross-sectional distribution of the number of foreign contacts across firms in Section 2.2, I would predict this marginal effect to be equal to .46%.²⁰ The cross-sectional estimate (.46%) and the time-series estimate (.48%) are close.

In addition, I can match the structural parameters in Equation (8) with the estimated parameters in Equation (12). The model predicts that the relative efficiency of the remote search and the direct search is measured by the parameter π . This relative efficiency is given by a simple ratio of parameters. When I use distance as a proxy for the search technology, $\beta_1 \approx \gamma\mu$ and $\beta_2 \approx \gamma\mu\pi$ so that $\pi \approx \beta_2/\beta_1$. Similarly when I use export growth, $\pi \approx \gamma_2/\gamma_1$. Intuitively, a firm is more likely to export to a close-by country; this direct search effect is captured by the parameter β_1 . But a firm is also more likely to export to a country that is close to countries where this firm already exports; this remote search effect is captured by the parameter β_2 . The ratio β_2/β_1 measures the relative efficiency of the remote and direct search technologies. The same argument goes for the coefficients γ_1 and γ_2 , where distance is replaced by the growth of exports towards countries where a firm tries to enter. Across the various specifications in columns (4), (5) and (6) of Table 2, I estimate π between 3.2% and 3.3% when using distance ($\pi \approx \beta_2/\beta_1$), and between 11.9% and 26.5% when using import growth ($\pi \approx \gamma_2/\gamma_1$). While lower with the distance proxy, with the export growth proxy, the estimated coefficient π is close to the 15.8% estimated from the cross-sectional distribution of the number of foreign markets across firms in Section 2.2.

To confirm the results above, I perform a series of robustness checks.²¹ For instance, using various lags of exports, I show that the data would not be consistent with a model where exporters enter countries in which they have acquired a distribution network, or sell to consumers that have a particular taste for their good, and where either the cost of a distribution network or tastes are spatially correlated.

Finally, using only information on the number of countries accessed, I structurally estimate²² the law of motion for the number of contacts in Equation (3),

$$M_{t+1} - M_t = \gamma\mu + \gamma\mu\pi M_t$$

²⁰Equation (3) in Section 1 gives $\Delta M_{i,t} = \gamma\mu\pi(1/\pi + M_{i,t})$. So adding one contact increases the growth in the number of contacts by $\gamma\mu\pi$. The average probability of entering a new country in the sample is $\Delta\bar{M} \approx 4.7\%$, the average number of contacts in the sample is $\bar{M} \approx 3.9$, and $1/\pi \approx 6.32$ from estimating Proposition 3 on the cross-sectional distribution of the number of foreign contacts. I predict that the increment in the probability of entering a new country stemming from moving from 3.9 to 4.9 contacts is given by $\gamma\mu\pi = \frac{\Delta\bar{M}}{(1/\pi + \bar{M})} \approx \frac{4.7\%}{6.32+3.9} \approx .46\%$.

²¹See the results in Tables 7 and 8 in Appendix B.

²²See the results in Table 9 in Appendix B.

Adding a series of controls, including on the growth rate of domestic sales of those firms to control for the growth trajectory a firm follows, does not affect those results substantially. A simple OLS estimation of the relationship above gives $\widehat{\gamma\mu\pi} = 0.165$ (.00040) and $\widehat{\gamma\mu} = 0.876$ (.0048). This implies $\pi \approx 0.188$, which is close to $\pi \approx 0.158$ estimated from the cross-sectional distribution of the number of foreign contacts across firms.

Conclusion

I have developed a theoretical model of the dynamic entry of firms into foreign markets. Firms can only export in countries where they have a contact. I assume that firms both directly search for foreign trading partners, and also use their existing network of contacts to remotely search for new partners. This dynamic model generates a stable distribution of exports across firms. The model makes precise predictions about the cross-sectional distribution of the number of foreign contacts, the cross-sectional distribution of the geographic distance from foreign contacts, and the dynamics of entry of individual firms into foreign markets. All theoretical predictions are supported by the data on firm level exports from France. Direct search is about 6 times more efficient than remote search.

This model and the empirical findings that support it suggest several extensions and generalizations. First, the emergence of a stable distribution of entrants into different foreign markets, and the fact that firms that export to more countries are less affected by geographic distance, may generate aggregate trade flows that follow the so called gravity equation. This may provide an explanation for the stable role that geographic distance plays in explaining bilateral trade flows. Second, I have only studied a simple symmetric case, and described its steady state properties. A large shock to this dynamic system would generate non trivial transitional dynamics. For example, a large disruption of trade linkages (e.g. wars or economic crises), or the rapid growth of a large country (e.g. China) may have a long lasting impact on the world geography of trade, since (re)building contacts is a lengthy and history dependent process. Third, whereas I have only sketched the welfare implications of a simple economic model that would support the proposed dynamics, the structure of the network lends itself to further welfare analysis. The robust predictions of the model regarding the geographic distribution of exports may allow for precise statements on the welfare gains from trade. I leave these questions for future research.

References

- ACEMOGLU, Daron, Vasco M. CARVALHO, Asuman OZDAGLAR and Alireza TAHBAZ-SALEHI. 2011. "The Network Origins of Business Fluctuations," MIT, *mimeo*.
- AHN, JaeBin, Amit K. KHANDELWAL and Shang-Jin WEI. 2011. "The Role of Intermediaries in Facilitating Trade," *Journal of International Economics*, 84(1): 73-85.
- ANTRÀS, Pol and Arnaud COSTINOT. 2011. "Intermediated Trade," *Quarterly Journal of Economics*, 126(3): 1319-1374.
- ALBORNOZ, Facundo, Hector F. CALVO PARDO, Gregory CORCOS, and Emanuel ORNELAS. 2010. "Sequential Exporting," CEPR Discussion Paper DP8103.
- ARMENTER, Roc and Miklós KOREN. 2009. "Economies of Scale and the Size of Exporters," Central European University, *mimeo*.
- ARMENTER, Roc and Miklós KOREN. 2010. "A Balls-and-Bins Model of Trade," Central European University, *mimeo*.
- ATALAY, Enghin, Ali HORTAÇSU, James W. ROBERTS, and Chad SYVERSON. 2010. "On the Network Structure of Production," University of Chicago, *mimeo*.
- BARABÁSI, Albert-László and Réka ALBERT. 1999. "Emergence of Scaling in Random Networks," *Science*, 286: 509-12.
- BERNARD, Andrew B., Jonathan EATON, J. Bradford JENSEN, and Samuel KORTUM. 2003. "Plants and Productivity in International Trade," *American Economic Review*, 93(4): 1268-90.
- BRAMOULLÉ, Yann and Brian W. ROGERS. 2010. "Diversity and Popularity in Social Networks," Northwestern University, *mimeo*.
- BURCHARDI, Konrad B. and Tarek A. HASSAN. 2010. "The Economic Impact of Social Ties: Evidence from German Reunification." Chicago Booth, *mimeo*.
- CASELLA, Alessandra and James E. RAUCH. 2002. "Anonymous Market and Group Ties in International Trade." *Journal of International Economics*, 58(1): 19-47.
- CHANEY, Thomas. 2008. "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4): 1707-21.
- COMBES, Pierre-Philippe, Miren LAFOURCADE and Thierry MAYER. 2005. "The Trade-Creating Effects of Business and Social Networks: Evidence from France," *Journal of International Economics*, 66(1):1-29.
- CURRARINI, Sergio, Matthew O. JACKSON and Paolo PIN. 2011. "Long-Run Integration in Social

- Networks,” Stanford University, *mimeo*.
- DEFEVER, Fabrice, Benedikt HEID and Mario LARCH. 2010. “Spatial Exporter Dynamic,” *mimeo*.
- EATON, Jonathan, Marcela ESLAVA, C.J. KRIZAN, Maurice KUGLER, and James TYBOUT. 2010. “A Search and Learning Model of Export Dynamics,” Penn State University, *mimeo*.
- EATON, Jonathan, Samuel KORTUM, and Francis KRAMARZ. 2011. “An Anatomy of International Trade: Evidence from French Firms,” *Econometrica*, 79(5): 1453-1498.
- ERDÖS, Paul and Alfréd RÉNYI. 1959. “On Random Graphs,” *Publicationes Mathematicae*, 6:290-7.
- FEENSTRA, Robert C., Robert E. LIPSEY, Haiyan DENG, Alyson C. MA, and Hengyong MO. 2004. “World Trade Flows: 1962-2000,” NBER WP 11040.
- GABAIX, Xavier. 1999. “Zipf Law for Cities: an Explanation,” *Quarterly Journal of Economics*, 114(3): 739-67.
- HIDALGO, César A., B. KLINGER, Albert-László BARABÁSI, and Ricardo HAUSMANN. 2007. “The Product Space Conditions the Development of Nations” *Science*, 317(5837): 482-487.
- HOLMES, Thomas J. 2011. “The Diffusion of Wal-mart and the Economies of Density” *Econometrica*, 79(1): 253–302.
- JACKSON, Matthew O., and Brian W. ROGERS. 2007. “Meeting Strangers and Friends of Friends: How Random Are Social Networks?” *American Economic Review*, 97(3): 890-915.
- JIA, Panle. 2008. “What Happens When Wal-mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry” *Econometrica*, 76(6): 1263–1316.
- KRUGMAN, Paul. 1980. “Scale Economies, Product Differentiation, and the Patterns of Trade,” *American Economic Review*, 70(5): 950-59.
- LUTTMER, Erzo G. J. 2007. “Selection, Growth, and the Size Distribution of Firms,” *Quarterly Journal of Economics*, 122(3): 1103-44.
- MAYER, Thierry, and Soledad ZIGNAGO. 2006. “Notes on CEPII’s Distances Measures,” *mimeo*.
- MORALES, Eduardo, Gloria SHEU, and Andrés ZAHLER. 2011. “Gravity and Extended Gravity: Estimating a Structural Model of Export Entry,” Harvard University, *mimeo*.
- MCPHERSON, Miller, Lynn SMITH-LOVIN, and James M. COOK. 2001. “Birds of a Feather: Homophily in Social Networks,” *Annual Review of Sociology*, 27: 414-44
- MELITZ, Marc J. 2003. “The Impact of Trade on Intra-Industry Reallocation and Aggregate Industry Productivity,” *Econometrica*, 71(6): 1695-1725.
- RAUCH, James E. 1999. “Networks Versus Markets in International Trade,” *Journal of Economic*

Literature, 48(1): 7-35.

RAUCH, James E. 2001. "Business and Social Networks in International Trade," *Journal of Economic Literature*, 39: 1177-1203.

RAUCH, James E. and Vitor TRINDADE. 2002. "Ethnic Chinese Networks in International Trade," *Journal of International Economics*, 84(1): 116-30.

ROSSI-HANSBERG, Esteban and Mark WRIGHT. 2007. "Establishment Size Dynamics in the Aggregate Economy," *American Economic Review*, 97(5): 1639-1666.

STEINDL, Josef. 1965. "Random Processes and the Growth of Firms," Charles Griffin, London.

APPENDIX

A Mathematical proofs

Lemma 1 *The density of consumers f_t is given by,*

$$f_t = \frac{1}{\pi} ((\delta + \gamma\mu\pi g)^{*t} - \delta)$$

where δ is the Dirac delta function and the exponent $*t$ stands for a function convoluted t times with itself. f_t admits a closed form solution in the special cases where g is a Gaussian or a Cauchy distribution.

Proof. First note that the integral on the right hand side of Equation (2) is a convolution product, so that the difference equation can be written in a compact form as,

$$f_{t+1} = \gamma\mu g + f_t + \gamma\mu\pi g * f_t$$

where $*$ stands for the convolution product of two functions. Taking a Fourier transform of this equation, and using the convolution theorem which states that the Fourier transform of the convolution of distributions is the product of their Fourier transforms, I get,

$$\hat{f}_{t+1} = \gamma\mu\hat{g} + \hat{f}_t + \gamma\mu\pi\hat{g}\hat{f}_t$$

with initial condition $\hat{f}_0 = 0$, where I denote by \hat{f} the Fourier transform of the function f . This first order linear recursive equation admits the following solution,

$$\hat{f}_t = \frac{1}{\pi} ((1 + \gamma\mu\pi\hat{g})^t - 1)$$

Taking the inverse Fourier transform of this equation, I get the proposed expression for f_t .

To derive a closed form solution for the special cases where g is a Gaussian or a Cauchy distribution, I manipulate this expression and get,

$$\begin{aligned} f_t &= \frac{1}{\pi} ((\delta + \gamma\mu\pi g)^{*t} - \delta) \\ &= \frac{1}{\pi} \sum_{s=1}^t (\gamma\mu\pi)^s \binom{t}{s} g^{*s} \end{aligned}$$

Note that the convolution of the p.d.f.'s of t random variables is the p.d.f of their sum. As the sum of t Gaussian (respectively Cauchy) distributed random variables is also a Gaussian (resp. Cauchy), I derive closed form solutions. If $g = \phi_{\sigma^2}$ where ϕ_{σ^2} is the p.d.f. of a Gaussian distribution with mean zero and variance σ^2 , then $g^{*s} = \phi_{s\sigma^2}$, and I get,

$$f_t = \frac{1}{\pi} \sum_{s=1}^t (\gamma\mu\pi)^s \binom{t}{s} \phi_{s\sigma^2}$$

If $g = \psi_\gamma$ where ψ_γ is the p.d.f. of a Cauchy distribution centered around zero and with scale parameter γ , then $g^{*s} = \psi_{s\gamma}$, and I get,

$$f_t = \frac{1}{\pi} \sum_{s=1}^t (\gamma\mu\pi)^s \binom{t}{s} \psi_{s\gamma}$$

■

Proposition 1 (reminded) *For short time intervals, the fraction of firms with fewer than m consumers is,*

$$F(m) = 1 - \left(\frac{1}{1 + \pi m} \right)^{\frac{1}{\mu\pi}}$$

Proof. At any time, in any location, and therefore in the union of any set of locations (any country), the fraction of firms with more than m contacts is the same. From the previous Lemma, integrating the density of contacts of a firm of age t , I get the number of contacts of a firm as a function of its age,

$$m_t = \frac{1}{\pi} ((1 + \gamma\mu\pi)^t - 1)$$

I invert the previous equation to get the age of a firm as a function of its number of contacts,

$$t(m) = \frac{\ln(1 + \pi m)}{\ln(1 + \gamma\mu\pi)}$$

where t only takes integer values.²³ The fraction of firms with more than m contacts, $1 - F(m)$, is the fraction of firms older than $t(m)$. Given the exponential growth rate of the population, this fraction is $(1 + \gamma)^{-t(m)}$. Using the above expression for $t(m)$, I get,

$$\begin{aligned} 1 - F(m) &= (1 + \gamma)^{-t(m)} \\ &= (1 + \pi m)^{-\ln(1+\gamma)/\ln(1+\gamma\mu\pi)} \end{aligned}$$

²³Note that I am not making any continuous approximation of the discrete model. The proposed formulas are exactly correct when t is an integer. Those formulas simply extrapolate to non integer values for t .

For short time intervals, which corresponds to γ small, the following first order approximation is exact,

$$\lim_{\gamma \rightarrow 0} \frac{\ln(1 + \gamma)}{\ln(1 + \gamma\mu\pi)} = \frac{1}{\mu\pi}$$

from which I derive the proposed expression,

$$F(m) = 1 - \left(\frac{1}{1 + \pi m} \right)^{\frac{1}{\mu\pi}}$$

■

Proposition 2 (reminded) *For short time intervals, the average (squared) distance from a firm's consumers, $\Delta(m)$, increases with the number of consumers m ,*

$$\Delta(m) = \left(1 + \frac{1}{\pi m} \right) \ln(1 + \pi m) \Delta_g$$

with $\Delta_g \equiv \int_{\mathbb{R}} x^2 g(x) dx$ the variance of the p.d.f. $g(\cdot)$.

Proof. From Proposition 1, I get an expression not only for the p.d.f. of the distribution of contacts, $g_t = f_t/m_t$, but more interestingly for its Fourier transform, $\hat{g}_t = \hat{f}_t/m_t$,

$$\hat{g}_t = \frac{(1 + \gamma\mu\pi\hat{g})^t - 1}{(1 + \gamma\mu\pi)^t - 1}$$

Note that if g_t is the p.d.f. of a random variable X_t , then its Fourier transform \hat{g}_t is closely related to $\varphi_{g_t}(\omega) = \mathbb{E}[e^{i\omega X_t}]$, the characteristic function of X_t ,

$$\hat{g}_t(\omega) = \int_{\mathbb{R}} g_t(x) e^{-i\omega x} dx = \mathbb{E}[e^{-i\omega X_t}] = \varphi_{g_t}(-\omega)$$

The various moments of g_t are then simply given by the various derivatives of \hat{g}_t evaluated at zero.

$$\begin{aligned} \hat{g}'_t &= \frac{\gamma\mu\pi t \hat{g}' (1 + \gamma\mu\pi\hat{g})^{t-1}}{(1 + \gamma\mu\pi)^t - 1} \\ \hat{g}''_t &= \gamma\mu\pi t \frac{\hat{g}'' (1 + \gamma\mu\pi)^{t-1} + \gamma\mu\pi (t-1) \hat{g}'^2 (1 + \gamma\mu\pi\hat{g})^{t-2}}{(1 + \gamma\mu\pi)^t - 1} \end{aligned}$$

Note that since the distribution g is symmetric about zero, its first moment is zero, $\hat{g}'(0) = 0$. The average (squared) distance of exports for a firm of age t , Δ_t , is simply the variance of the p.d.f. g_t , given by the second derivative of \hat{g}_t evaluated at zero.

$$\Delta_t \equiv \int_{\mathbb{R}} x^2 g_t(x) dx = \mathbb{E}[X_t^2] = \hat{g}''_t(0) = \frac{\gamma\mu\pi t (1 + \gamma\mu\pi)^{t-1}}{(1 + \gamma\mu\pi)^t - 1} \Delta_g$$

with $\Delta_g \equiv \int_{\mathbb{R}} x^2 g(x) dx$ the variance of the p.d.f. $g(\cdot)$. Using the expression that relates a firm's age to the number of its contacts from the proof of the previous proposition, I get,

$$\begin{aligned} t(m) &= \frac{\ln(1 + \pi m)}{\ln(1 + \gamma \mu \pi)} \\ (1 + \gamma \mu \pi)^{t(m)} - 1 &= \pi m \\ (1 + \gamma \mu \pi)^{t(m)-1} &= \frac{1 + \pi m}{1 + \gamma \mu \pi} \end{aligned}$$

plugging those expressions into the expression for $\Delta_{t(m)} = \Delta(m)$, I get,

$$\Delta(m) = \frac{\gamma \mu \pi}{(1 + \gamma \mu \pi) \ln(1 + \gamma \mu \pi)} \left(1 + \frac{1}{\pi m}\right) \ln(1 + \pi m) \Delta_g$$

For short time intervals, which corresponds to γ small, by l'Hôpital's rule, the following first order approximation is exact,

$$\lim_{\gamma \rightarrow 0} \frac{\gamma \mu \pi}{(1 + \gamma \mu \pi) \ln(1 + \gamma \mu \pi)} = 1$$

from which I derive the proposed expression,

$$\Delta(m) = \left(1 + \frac{1}{\pi M m}\right) \ln(1 + \pi m) \Delta_g$$

■

Proposition 3 (reminded) *If there are N equal size distinct foreign countries, and the distribution of the number of consumers across firms is given by Proposition 1, then the fraction of firms that export to exactly M distinct countries is,*

$$\phi(M|N) = \sum_{m=M}^{+\infty} \frac{1}{\mu} \left(\frac{1}{1 + \pi m}\right)^{\frac{1}{\mu\pi}+1} n! S_2(m, M) \binom{N}{M} \left(\frac{1}{N}\right)^m$$

with $S_2(m, M)$ the Stirling number of the second kind.

Proof. There are N distinct countries, each populated by the same number of firms. These firms are connected to one another by the network described in Section 1. The distribution of the number of contacts M is therefore given by the c.d.f. $F(M)$ in Proposition 1. Assume for simplicity that physical geography does not matter, so that any contact is equally likely to be located in any of the N countries, and that the location of any two contacts are independent from one another. In other words, there is a probability $1/N$ that a contact is located in a given country, and those probabilities are independent across contacts.²⁴

²⁴In the model, contacts are spatially correlated, so that some configurations of contacts are more likely than others. Not accounting for this spatial correlation may induce a downward bias in the estimated number of contacts. Despite this, I keep the assumption of equi-probable configurations for tractability. Numerical simulations suggest that this bias is not systematically different for firms with many or few contacts, so that the estimation of the parameters does not suffer much from this simplifying assumption.

Consider now the following question: what is the probability, $\Pr(M|m)$, that a firm with m contacts has contacts in M distinct countries?

For a given number of contacts m , there are N^m equi-probable distinct ways of distributing m contacts into the N countries of the world, so that each configuration is realized with a probability $(\frac{1}{N})^m$. There are $\binom{N}{M}$ different ways of choosing M countries out of the total of N countries. Call $G(m, M)$ the number of distinct ways to assign each of the m contacts into M countries. The probability that a firm with m contacts has contacts in M distinct countries is then simply,

$$\Pr(M|m) = \binom{N}{M} \left(\frac{1}{N}\right)^m G(m, M)$$

The number $G(m, M)$ is defined recursively as follows. There are two mutually exclusive cases. In the first case, the first $m - 1$ contacts are assigned to only $M - 1$ countries. In that case, the last m^{th} contact must necessarily be assigned to the M^{th} country. There are M equi-probable such cases, one for each M^{th} missing last country. There are $G(m - 1, M - 1)$ ways to assign $M - 1$ countries to $m - 1$ contacts, M candidate countries that can be missing for the last contact. There are therefore $M G(m - 1, M - 1)$ distinct ways of assigning M countries to m contact in that first case. In the second case, the first $m - 1$ contacts are assigned to M countries. We can then assign the last contact to any one of the M countries. There are therefore $M G(m - 1, M)$ distinct ways of assigning M countries to m contacts in that second case. The number $G(m, M)$ is defined recursively as,

$$G(m, M) = M [G(m - 1, M - 1) + G(m - 1, M)]$$

with the initial conditions $G(m, m) = m!$ and $G(m, 1) = 1$. Noting that the known Stirling number of the second kind, $S_2(m, M)$, is defined recursively in a similar fashion,

$$S_2(m, M) = S_2(m - 1, M - 1) + M S_2(m - 1, M)$$

with the initial conditions $S_2(m, m) = 1$ and $S_2(m, 1) = 1$, I get the following relationship between the numbers $G(m, M)$ and $S_2(m, M)$,

$$G(m, M) = M! S_2(m, M)$$

I can now answer the question of interest: given that the number of contacts, m , is distributed according to the distribution $F(m)$, and that there are N equal sized countries, what is the distribution of the number of countries, M , accessed by different firms?

Obviously, a firm that has m contacts can export to at most m different countries. But for any $m > 1$, the probability that two different contacts fall into the same country is positive, and increases with m . In other words, among firms with contacts in M distinct countries, there are firms with M , $M + 1$, $M + 2$, \dots contacts.

The fraction of firms that have exactly m contacts is simply given by the p.d.f. $f(m)$,²⁵ associated with the c.d.f. $F(m)$ defined in Proposition 1. Given the distribution of the number of countries reached by a firm with m contacts, $\Pr(M|m)$, the fraction of firms that have contacts in exactly M distinct countries, is given by the proposed expression,

$$\begin{aligned}\phi(M|N) &= \sum_{m=M}^{+\infty} f(m) \Pr(M|m) \\ &= \sum_{m=M}^{+\infty} \frac{1}{\mu} \left(\frac{1}{1 + \pi m} \right)^{\frac{1}{\mu\pi} + 1} n! S_2(m, M) \binom{N}{M} \left(\frac{1}{N} \right)^m\end{aligned}$$

■

B Data

In this section, I describe the source of the data. I provide some descriptive statistics in addition to what is presented in the main body of the paper. And I perform a series of robustness checks.

B.1 Data sources

Firm level export data: The data on firm level exports come from the French customs, and are described in greater detail in Eaton, Kortum and Kramarz (2011). Until 1992, all shipments crossing the French border are reported, either by the owner of the (exporting) firm, or by authorized customs commissioners. Information about the identity of the exporting firm, the value of the shipment, the industrial sector, and the destination country is recorded. This information is then aggregated over a year. I use data on all French exporters (including non manufacturing firms).²⁶ A data point is therefore a firm, year, destination country and value of exports (in French Francs) vector. Since I am primarily interested in the extensive margin of exports, I do not use information on the value of exports.

²⁵For simplicity, I use the approximation $f(m) \approx F(m+1) - F(m)$.

²⁶Restricting the sample to manufacturing firms does not alter the results significantly.

Table 4: Industrial sectors

Sector	N100 Industries	Firm-Years
Agriculture	0 – 3	302202
Mining	4 – 14	201983
Construction	15 – 15	298906
Manufacturers	16 – 56	10299279
Transportation	68 – 75	161195
Wholesale	57 – 59	5171218
Retail	60 – 64	1556536
F.I.R.E	78 – 81 , 88–89	148938
Services	65 – 67, 77, 82–87	905370
Public Admin. + Other	90 – 99	13390

In addition, the customs data are matched with balance sheet information collected by the French fiscal authorities. All firms with a turnover of 1,000,000 French Francs in services, or 3,000,000 French Francs in manufacturing are mandated to report this information. Virtually all exporters are included in this data set. In some robustness checks (Table 9), I use information on annual sales, employment, and capital expenditure.

Finally, I use information on the primary 2-digit industrial sector of a firm. Table 4 reports the list of 2-digit sectors, as well as the distribution of all exporters in those sectors.

For the main regressions of interest, I restrict my sample of firms to exporters only.

Distance data: I use data on bilateral distances between countries collected and constructed by the CEPII. The distance between two countries is calculated as a weighted arithmetic average of the geodesic distances between the main cities in these countries, where population weights are used. Data on the location of the main cities in each country (latitude and longitude), as well as the population of those main cities are used to compute those distances. The construction of the data is described in further details by Mayer and Zignago (2006).

Country size data: I use as a measure of a country’s size its nominal *GDP* (in US\$) in the current year. The data are collected from the Penn World Tables and are described in further detail at <http://pwt.econ.upenn.edu/>.

Bilateral trade flows: To proxy for the intensity of firm level contacts between countries other than France, I use data on bilateral trade flows between countries. The data correspond to the nominal value (in US\$) of aggregate trade flows between country pairs. The data are collected

from the NBER, and are described in further detail in Feenstra et al. (2004).

B.2 Robustness checks

Cross-section: $\phi(M|N)$ versus M

To check the robustness of the results presented in Section 2.2, I replicate the estimations in columns (1)-(3) in Table 1 for each year, from 1986 to 1992. That is, I estimate separately for each year the prediction from Proposition 1 regarding the cross-sectional distribution of the number of foreign contacts using Maximum Likelihood after controlling for the fact that I observe countries and not contacts. In addition, I estimate the parameters using Maximum Likelihood and Non Linear Least Squares under the simplifying that firms have a single contact per country where they export.

The results are presented in Table 5. All estimated parameters are statistically different from zero at the 1% level of significance. Except for the year 1986 where the estimated π is somewhat lower in the NLLS estimation, the coefficients are virtually identical across years, and do not differ much across the different estimation procedures. This suggests first that the results are robust, and second that the system of French exporters is in a steady state over the period considered.

Cross-section: $\Delta(M)$ versus M

To check the robustness of the results presented in Section 2.3, I estimate the relevant parameter Δ_g using different empirical measures of $\Delta(M)$ and different years.

The results are presented in Table 6. For each year between 1986 and 1992, I separately estimate the parameter Δ_g . To do so, I use the formula for $\Delta(M)$ derived in Proposition 2, imposing the parameter π which is estimated on the cross-sectional distribution of the number of foreign contacts in the top panel of Table 5. I estimate Δ_g using non linear least squares, weighting each observation by the precision of the empirical estimate of $\Delta(M)$. I use two alternative empirical measures for $\Delta(M)$.

The top panel of Table 6 uses the same empirical measure of the second moment of the distance from a firm's export destinations, $\Delta(M)$, as in Section 2.3, where I correct for differences in *GDP* across countries. The formula for calculating the empirical counterpart of $\Delta(M)$ is,

$$\widehat{\Delta(M)} = \frac{\sum_{i \in E(M),c} (Distance_{Fr,c})^2 \left(\frac{1}{GDP_c} \right) \mathbb{I}\{export_{i,c} > 0\}}{\sum_{i \in E(M),c} \left(\frac{1}{GDP_c} \right) \mathbb{I}\{export_{i,c} > 0\}}$$

Table 5: Empirical fit of Proposition 3, robustness checks

Year	(1986)	(1987)	(1988)	(1989)	(1990)	(1991)	(1992)
Maximum Likelihood Estimation with Correction of $\phi(M N = 210)$:							
π	0.165 (.0018)	0.164 (.0017)	0.156 (.0018)	0.156 (.0017)	0.153 (.0017)	0.153 (.0017)	0.158 (.0017)
μ	2.479 (.0105)	2.460 (.0101)	2.432 (.0104)	2.442 (.0102)	2.433 (.0103)	2.399 (.0099)	2.384 (.0098)
log (likelihood)	-285,711	-297,839	-268,454	-280,670	-270,464	-277,844	-282,098
Maximum Likelihood Estimation of $f(M)$:							
π	0.154 (.0017)	0.153 (.0017)	0.146 (.0017)	0.146 (.0017)	0.143 (.0017)	0.144 (.0019)	0.148 (.0017)
μ	2.491 (.0105)	2.473 (.0102)	2.443 (.0104)	2.453 (.0102)	2.443 (.0103)	2.409 (.0105)	2.394 (.0098)
log (likelihood)	-286653	-297774	-268388	-280602	-270395	-277770	-282022
Non Linear Least Squares Estimation of $f(M)$:							
π	0.102 (.016)	0.125 (.018)	0.134 (.025)	0.129 (.023)	0.133 (.025)	0.131 (.022)	0.130 (.021)
μ	3.574 (.386)	3.152 (.332)	2.916 (.381)	2.983 (.377)	2.900 (.382)	2.934 (.354)	2.970 (.349)
Adj. R^2	0.975	0.980	0.971	0.972	0.971	0.976	0.976

Notes: This table shows the results of the Maximum Likelihood and Non Linear Least Squares estimations of the parameters (μ, π) in Proposition 1 for French exporters in all years from 1986 to 1992. The top panel estimates by maximum likelihood the p.d.f. $\phi(M|N = 210)$ using the statistical correction proposed in Proposition 3 for the fact that the number of countries, M , and not contacts, m , is observed. The middle panel estimates by maximum likelihood the p.d.f. $f(M)$ under the simplifying assumption that the number of contacts of a firm and the number of countries it exports to are the same, $m = M$. The bottom panel uses the log of the fraction of firms that export to M markets as the dependent variable, and estimates (μ, π) using a non linear least squares estimation of Proposition 1, under the same assumption $m = M$. Standard errors are reported in parentheses. All coefficients on this table are statistically different from zero at the 1% level of significance.

Table 6: Empirical fit of Proposition 2, robustness checks

Year	(1986)	(1987)	(1988)	(1989)	(1990)	(1991)	(1992)
$\Delta(M)$ corrected for <i>GDP</i> differences:							
Δ_g	14.434	14.325	14.602	14.850	15.01	14.819	14.860
	(.124)	(.121)	(.118)	(.117)	(.103)	(.085)	(.109)
Adj. R^2	0.859	0.888	0.872	0.881	0.887	0.924	0.871
$\Delta(M)$ not corrected for <i>GDP</i> differences:							
Δ_g	14.437	14.178	14.412	14.730	14.439	14.346	14.384
	(.292)	(.318)	(.335)	(.311)	(.360)	(.401)	(.373)
Adj. R^2	0.820	0.792	0.784	0.801	0.766	0.736	0.753

Notes: This table shows the non linear least square estimate of Δ_g from Proposition 2, imposing the parameter π estimated by MLE in the top panel of Table 5. The estimation is run separately for each year from 1986 to 1992. The top panel corrects the empirical measure of $\Delta(M)$ for differences in *GDP* across countries, while the bottom panel does not. Each point weighted by the square root of the number of observations used to compute $\Delta(M)$. Standard errors are in parentheses. All coefficients on this table are statistically different from zero at the 1% level of significance.

where $E(M)$ is the set of firms that export to M countries.

In the second panel of Table 6, I do not correct for differences in *GDP* across countries. The formula for calculating $\Delta(M)$ is as follows,

$$\widehat{\Delta(M)} = \frac{\sum_{i \in E(M), c} (Distance_{Fr, c})^2 \mathbb{I}\{export_{i, c} > 0\}}{\sum_{i \in E(M), c} \mathbb{I}\{export_{i, c} > 0\}}$$

where $E(M)$ is the set of firms that export to M countries.

This correction does not affect the estimated Δ_g substantially. The statistical significance of the estimated Δ_g is reduced, and the R^2 goes down from about 90% to 75% when I do not control for differences in *GDP*. But all the coefficients remain highly significant (at the 1% confidence level), and the coefficients themselves do not differ much across both measures. The estimated coefficients are very stable over time. This suggests that the results presented in Section 2.3 are robust. Moreover, given that I only allow for a single degree of freedom when estimating this relationship, Proposition 2 finds a remarkably strong support in the data.

Time-series: PROBIT regression

To check the robustness of the results presented in Section 2.4, I run a Probit regression of

different specifications of the following equation,

$$\begin{aligned}
\mathbb{I}\{export_{i,c,t} > 0\} = & \alpha \times \{\text{N. contacts}_{i,t-1}\} \\
& + \beta_1 \times \sum_{c' \in C_{i,t-1}} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t-1}} + \beta_2 \times \sum_{c' \in C_{i,t-2}} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t-2}} + \dots \\
& + \gamma_1 \times \frac{1}{|C_{i,t-1}|} \sum_{c' \in C_{i,t-1}} \ln Dist_{c',c} + \gamma_2 \times \frac{1}{|C_{i,t-2}|} \sum_{c' \in C_{i,t-2}} \ln Dist_{c',c} + \dots \\
& + \mu \times \ln Dist_{Fr,c} + \nu \times \frac{1}{|\{c' : c' \neq Fr\}|} \sum_{c' \neq Fr} \ln Dist_{c',c} \\
& + \delta_1 \times \mathbb{I}\{export_{i,c,t-1} > 0\} + \delta_2 \times \mathbb{I}\{export_{i,c,t-2} > 0\} + \dots \\
& + \text{Controls} + \epsilon_{i,c,t} \quad (13)
\end{aligned}$$

where $C_{i,t-k}$ is the set of countries where firm i exports at time $t - k$.

Table 7 shows the results of these Probit regressions, and Table 8 the corresponding marginal effects. With a few exceptions, all coefficients are statistically significant (at the 1% confidence level), and of the expected signs.

The results of these regressions allow me to reject an alternative model. In this alternative model, firms either have to acquire a local distribution network to enter foreign countries, or sell goods to consumers with locally differentiated tastes. Moreover, the cost of expanding a distribution network or local tastes are spatially correlated. In such an alternative model, a firm would be more likely to enter at time t a market c that is geographically close to, or that trade a lot with any market c' in which this firms exported at time $t - 1$. However, in such an alternative model, conditional on exporting to countries that are either close or trades a lot with c at time $t - 1$, the increment in the probability of entering c given that a firm exports to countries that are close or trade a lot with c at times $t - 2, t - 3, \dots$ should decrease rapidly. In my model on the other hand, any year a firm exports to country c' , and conditional on still not exporting to c , it has the same probability of learning about contacts in c . So controlling for the export status in c over longer and longer lags, the increment in probability of entering c brought about by the fact that a firm exports to countries that are close or trade a lot with c should not systematically decrease over longer and longer lags.

As can be seen in Columns (2)-(5), there is no systematic tendency for the coefficients β_1, \dots, β_5 and $\gamma_1, \dots, \gamma_5$ on these lags to fall. In Column (5) where I use the maximum number of five lags allowed by the data, the coefficients on the lags of export growth, β 's, fall and then increase, while the coefficients on the lags of distance, γ 's, increase, except for the fifth lag that is insignificant.

Table 7: Time-series of exports, robustness checks (PROBIT)

	Dependent Variable: $\mathbb{I}\{export_{i,c,t} > 0\}$				
	(1)	(2)	(3)	(4)	(5)
N. contacts _{<i>i,t-1</i>}	.0343*** (.0003)	.0293*** (.0003)	.0275*** (.0003)	.0261*** (.0003)	.0259*** (.0004)
$\sum_{c' \in C_{i,t-1}} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t-1}}$.0355*** (.0028)	.0406*** (.0036)	.0560*** (.0043)	.0363*** (.0053)	.0359*** (.0067)
$\sum_{c' \in C_{i,t-2}} \frac{\Delta Exports_{c',c,t-1}}{Exports_{c',c,t-2}}$.0083*** (.0011)	.0001 (.0015)	-.0070*** (.0019)	-.0115*** (.0027)
$\sum_{c' \in C_{i,t-3}} \frac{\Delta Exports_{c',c,t-2}}{Exports_{c',c,t-3}}$.0198*** (.0015)	.0111*** (.0019)	.0059*** (.0027)
$\sum_{c' \in C_{i,t-4}} \frac{\Delta Exports_{c',c,t-3}}{Exports_{c',c,t-4}}$.0237*** (.0018)	.0183*** (.0024)
$\sum_{c' \in C_{i,t-5}} \frac{\Delta Exports_{c',c,t-4}}{Exports_{c',c,t-5}}$.0832*** (.0025)
$\frac{\sum_{c' \in C_{i,t-1}} \ln Dist_{c',c}}{ C_{i,t-1} }$	-.1477*** (.0032)	-.955*** (.0041)	-.0720*** (.0053)	-.0509*** (.0067)	-.0372*** (.00086)
$\frac{\sum_{c' \in C_{i,t-2}} \ln Dist_{c',c}}{ C_{i,t-2} }$		-.0740*** (.0039)	-.0582*** (.0054)	-.0470*** (.0067)	-.0432*** (.00089)
$\frac{\sum_{c' \in C_{i,t-3}} \ln Dist_{c',c}}{ C_{i,t-3} }$			-.0491*** (.0050)	-.0458*** (.0067)	-.0459*** (.0089)
$\frac{\sum_{c' \in C_{i,t-4}} \ln Dist_{c',c}}{ C_{i,t-4} }$				-.0437*** (.0062)	-.0364*** (.0085)
$\frac{\sum_{c' \in C_{i,t-5}} \ln Dist_{c',c}}{ C_{i,t-5} }$.202*** (.0083)
$\ln Dist_{Fr,c}$	-.3806*** (.0031)	-.2923*** (.0032)	-.2592*** (.0037)	-.2449*** (.0045)	-.2127*** (.0055)
$\frac{1}{ \{c': c' \neq Fr\} } \sum_{c' \neq Fr} \ln Dist_{c',c}$.4397*** (.0075)	.3621*** (.0078)	.3387*** (.0091)	.3393*** (.0113)	.2714*** (.0140)
$\ln GDP_{c,t}$.1296*** (.0008)	.1154*** (.0008)	.1124*** (.0009)	.1114*** (.0010)	.1146*** (.0013)
Control for: $\mathbb{I}\{export_{i,c,t-k} > 0\}$ and $\ln \left(\frac{\text{Domestic Sales}_t}{\text{Domestic Sales}_{t=k}} \right)$ $\sum_{c' \neq Fr} \frac{\Delta Exports_{c',c,t-k}}{Exports_{c',c,t-k}}$	<i>k</i> = 1	<i>k</i> = 1, 2	<i>k</i> = 1, ..., 3	<i>k</i> = 1, ...4	<i>k</i> = 1, ...5
Constant	-3.244*** (.0543)	-3.0227*** (.0544)	-3.0218*** (.0616)	-3.0942*** (.0748)	-2.7457 (.0915)
N. obs	16,565,725	12,255,332	8,898,165	5,950,983	3,689,058
N. clusters	34,588	32,589	29,745	23,149	20,488
Pseudo- <i>R</i> ²	.5608	.46935	.6073	.6136	.6171

Notes: This table shows the results of the PROBIT estimation of Equation (13) for a panel of all French exporters between 1986 and 1992. See the description of the variables in Section 2.4 on page 25. Standards errors are clustered at the firm level. *, **, and *** mean statistically significant at the 10, 5 and 1% levels.

Table 8: Time-series of exports, robustness checks (PROBIT: marginal effects)

Dep. Var.: $\mathbb{I}\{export_{i,c,t} > 0\}$	dy/dx				
	(1)	(2)	(3)	(4)	(5)
N. contacts $_{i,t-1}$.0016*** (.0000)	.0014*** (.0000)	.0013*** (.0000)	.0012*** (.0000)	.0012*** (.0000)
$\sum_{c' \in C_{i,t-1}} \frac{\Delta Exports_{c',c,t}}{Exports_{c',c,t-1}}$.0017*** (.0001)	.0019*** (.0002)	.0026*** (.0002)	.0016*** (.0002)	.0016*** (.0003)
$\sum_{c' \in C_{i,t-2}} \frac{\Delta Exports_{c',c,t-1}}{Exports_{c',c,t-2}}$.0004*** (.0001)	0 (.0001)	-.0003*** (.0001)	-.0005*** (.0001)
$\sum_{c' \in C_{i,t-3}} \frac{\Delta Exports_{c',c,t-2}}{Exports_{c',c,t-3}}$.0009*** (.0001)	.0005*** (.0001)	.0003*** (.0001)
$\sum_{c' \in C_{i,t-4}} \frac{\Delta Exports_{c',c,t-3}}{Exports_{c',c,t-4}}$.0011*** (.0001)	.0008*** (.0001)
$\sum_{c' \in C_{i,t-5}} \frac{\Delta Exports_{c',c,t-4}}{Exports_{c',c,t-5}}$.0038*** (.0004)
$\frac{\sum_{c' \in C_{i,t-1}} \ln Dist_{c',c}}{ C_{i,t-1} }$	-.0070*** (.0002)	-.0045*** (.0002)	-.0033*** (.0002)	-.0023*** (.0003)	-.0017*** (.0004)
$\frac{\sum_{c' \in C_{i,t-2}} \ln Dist_{c',c}}{ C_{i,t-2} }$		-.0035*** (.0002)	-.0027*** (.0002)	-.0021*** (.0003)	-.0020*** (.0004)
$\frac{\sum_{c' \in C_{i,t-3}} \ln Dist_{c',c}}{ C_{i,t-3} }$			-.0023*** (.0002)	-.0021*** (.0003)	-.0021*** (.0004)
$\frac{\sum_{c' \in C_{i,t-4}} \ln Dist_{c',c}}{ C_{i,t-4} }$				-.0020*** (.0003)	-.0017*** (.0004)
$\frac{\sum_{c' \in C_{i,t-5}} \ln Dist_{c',c}}{ C_{i,t-5} }$.0009*** (.0001)
$\ln Dist_{Fr,c}$	-.0181*** (.0001)	-.0136*** (.0001)	-.0120*** (.0002)	-.0110*** (.0002)	-.0097*** (.0003)
$\frac{1}{ \{c': c' \neq Fr\} } \sum_{c' \neq Fr} \ln Dist_{c',c}$.0209*** (.0003)	.0169*** (.0004)	.0156*** (.0004)	.0153*** (.0005)	.0124*** (.0006)
$\ln GDP_{c,t}$.0062*** (.0000)	.0054*** (.0000)	.0052*** (.0000)	.0050*** (.0000)	.0052*** (.0001)
Control for: $\mathbb{I}\{export_{i,c,t-k} > 0\}$ and $\ln \left(\frac{\text{Domestic Sales}_t}{\text{Domestic Sales}_{t=k}} \right)$ $\sum_{c' \neq Fr} \frac{\Delta Exports_{c',c,t-k}}{Exports_{c',c,t-k}}$	$k = 1$	$k = 1, 2$	$k = 1, \dots, 3$	$k = 1, \dots, 4$	$k = 1, \dots, 5$

Notes: This table shows the marginal effects for the PROBIT estimation of Equation (13) presented in Table 7. The marginal effect is calculated as dy/dx at the average value of each x in the sample. dy/dx is for a discrete change from 0 to 1 when x is a dummy variable. Standards errors are clustered at the firm level. *, **, and *** mean statistically different from zero at the 10, 5 and 1% level of significance.

It must be noted however that there is a lot of turnover in the extensive margin of trade at the firm level, so that the systematic entry and exit of firms into different markets may obscure some of these results.

Time-series: number of export destinations

The theoretical prediction from the model regarding the law of motion of the number of foreign contacts is given in Equation (3). To structurally test this prediction, I estimate by Ordinary Least Squares the following equation,

$$(\text{Number of new contacts})_{i,t+1} = \alpha + \beta M_{i,t} + \text{Controls}_{i,t} + \epsilon_{i,t} \quad (14)$$

where the dependent variable is the number of new countries entered by firm i between year t and $t + 1$, $M_{i,t}$ is the number of countries where firm i exports at time t , and $\epsilon_{i,t}$ is a normally distributed error term. Since I do not allow for the death of contacts in the theoretical model, I focus my empirical analysis on the creation of new contacts. The estimation of Equation (14) allows me to recover the structural parameter $\pi = \frac{\beta}{\alpha}$.

The results of this estimation are presented in Table 9. In the different specifications, I control for various measures of a firm's growth trajectory. The idea is that firms that grow on the domestic market may as well expand abroad, for reasons that are orthogonal to my model of network formation. I control for different combinations and measures of domestic sales growth, domestic employment growth, and domestic investment growth at the firm level.

Firms whose domestic sales are increasing are more likely to enter new foreign markets. Employment growth has some limited but non robust positive impact on the entry into new foreign markets, whereas investment growth does not seem to have any impact on the entry into foreign markets. If anything, investment growth deters entry into foreign markets.

The estimation of the various specifications of Equation (14) give an estimate for π that ranges between 0.5 (column (1) without any controls) to 0.176 (column (7)). The specification that is characterized by the most significant combination of controls, column (6), gives an estimate of 0.182. Given that an entirely different set of data is used, this is surprisingly close to $\pi \approx 0.158$ estimated from the cross-sectional distribution of the number of foreign contacts.

These findings suggest that my model of network formation is able to precisely identify the link between the time-series and the cross-section of individual exporters' entry into foreign markets.

Table 9: Time-series of exports, structural estimation of Proposition 1

	Dependent variable: number of new foreign markets entered by firm i between t and $t + 1$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$M_{i,t}$.182*** (0.00027)	.165*** (0.00040)	.166*** (0.00040)	.167*** (0.00040)	.167*** (0.00040)	.165*** (0.00040)	.165*** (0.00042)	.167*** (0.00040)	.165*** (0.00042)
$Sales_{i,t+1} - Sales_{i,t}$		8.37E-8*** (7.52E-9)				1.70E-7*** (1.06E-8)			
$\frac{Sales_{i,t+1}}{Sales_{i,t}}$			4.95E-6** (1.81E-6)				.000111*** (1.81E-5)		
$Sales_{i,t} - Sales_{i,t-1}$				3.47E-8*** (5.85E-9)				3.65E-8*** (8.25E-9)	
$\frac{Sales_{i,t}}{Sales_{i,t-1}}$					4.87E-6** (1.69E-6)				5.34E-5** (1.68E-5)
$Emp_{i,t+1} - Emp_{i,t}$						4.71E-5*** (8.60E-6)			
$\frac{Emp_{i,t+1}}{Emp_{i,t}}$							-3.30E-6 (2.21E-5)		
$Emp_{i,t} - Emp_{i,t-1}$								5.31E-6 (7.59E-6)	-1.35E-5 (2.06E-5)
$\frac{Emp_{i,t}}{Emp_{i,t-1}}$									
$Inv_{i,t+1} - Inv_{i,t}$						-6.67E-7*** (3.67E-8)			
$\frac{Inv_{i,t+1}}{Inv_{i,t}}$							-6.87E-9 (7.32E-7)		
$Inv_{i,t} - Inv_{i,t-1}$								-4.15E-8 (3.50E-8)	
$\frac{Inv_{i,t}}{Inv_{i,t-1}}$									2.54E-6*** (6.80E-7)
Constant	.360*** (0.0022)	.876*** (0.0048)	.874*** (0.0048)	.675*** (0.0048)	.674*** (0.0048)	.876*** (0.0048)	.908*** (0.0053)	.675*** (0.0048)	.708*** (0.00512)
N. obs.	743916	292773	291726	263492	262472	292773	258336	263492	232501
R^2	0.38	0.37	0.37	0.40	0.40	0.37	0.37	0.40	0.40

Notes: This table shows the results of the OLS estimation of different specifications of Equation (14) for French exporters over the period 1986 to 1992. The dependent variable is the number of new foreign markets entered by firm i between year t and $t + 1$. $M_{i,t}$ is the number of foreign markets where firm i exports at time t . $Sales_{i,t}$ is aggregate domestic sales (in French Francs, in France) of firm i at time t . $Emp_{i,t}$ is the total number of employees of firm i at time t . $Inv_{i,t}$ is the total capital expenditure (in French Francs) of firm i at time t . Standards errors are clustered at the firm level. *, **, and *** mean statistically different from zero at the 10, 5 and 1% level of significance.